

Delineating the field of language evolution research

A quantitative analysis of peer-review patterns at the Joint Conference on Language Evolution (JCoLE 2022)

Stefan Hartmann¹, Sławomir Wacewicz², Andrea Ravignani^{3,4}, Daria Valente⁵, Evelina Daniela Rodrigues⁶, Rie Asano⁷, and Yannick Jadoul^{3,4,8}

¹ Heinrich-Heine-Universität Düsseldorf | ² Nicolaus Copernicus University | ³ Max Planck Institute for Psycholinguistics | ⁴ Sapienza University of Rome | ⁵ University of Turin | ⁶ ISPA–Instituto Universitário | ⁷ Teikyo University | ⁸ Vrije Universiteit Brussel

Research on language evolution is an established subject area yet permeated by terminological controversies about which topics should be considered pertinent to the field and which not. By consequence, scholars focusing on language evolution struggle in providing precise demarcations of the discipline, where even the very central notions of evolution and language are elusive. We aimed at providing a data-driven characterisation of language evolution as a field of research by relying on quantitative analysis of data drawn from 697 reviews on 255 submissions from the Joint Conference on Language Evolution 2022 (Kanazawa, Japan). Our results delineate a field characterized by a core of main research topics such as iconicity, sign language, multimodality. Despite being explored within the framework of language evolution research, only very recently these topics became popular in linguistics. As a result, language evolution has the potential to emerge as a forefront of linguistic research, bringing innovation to the study of language. We also see the emergence of more recent topics like rhythm, music, and vocal learning. Furthermore, the community identifies cognitive science, primatology, archaeology, palaeoanthropology, and genetics as key areas, encouraging empirical rather than theoretical work. With new themes, models, and methodologies emerging, our results depict an intrinsically multidisciplinary and evolving research field, likely adapting as language itself.

Keywords: conference on language evolution, language, language evolution, research topics, controversies

1. Introduction

Research into the origins of language is distinguished by its rich intellectual history: long considered to be a domain of armchair philosophizing (Gould & Lewontin, 1979; Lewontin, 1998), it has now transformed into a qualitatively different enterprise, commonly referred to as *language evolution*. The perception of this new field of language evolution has consolidated in terms of institutional indicators, such as flagship publications, research centres, funded projects, and in particular conferences and journals (Dediu & de Boer, 2016), to yield a program that is institutionally as well as theoretically progressive (Żywicznyński, 2018). At the same time, it has been surrounded by terminological controversies about its scope: what research does or should count as “language evolution research”?

Not surprisingly, establishing or stipulating the scope of language evolution from a top-down perspective has proven to be difficult. On a general level, the two central notions of “evolution” (Haspelmath, 2016) and “language” (Waciewicz et al., 2020) have turned out to be very elusive. More specific controversies related, for example, to whether the term *language evolution* should be used exclusively to refer to the emergence of human linguistic capacities, or whether the traceable historical changes of human languages can also be legitimately referred to with the term *evolution* (Haspelmath, 2016; Mendivil-Giró, 2019). Gong et al. (2014), for instance, take issue with the tendency of pre-2014 iterations of the International Conference on Language Evolution (Evolang) to disregard contributions from the field of historical linguistics, thus hampering a potentially fruitful cross-fertilisation between the study of language evolution and language change (also see Hartmann, 2020). This case in point also illustrates that conferences – far from serving only to disseminate research – fulfill a more active role as tools for gate-keeping and policy-making: in addition to reflecting the field, they also co-create it by setting the trends as well as boundaries.

In sum, the practitioners of language evolution have had little success in providing accurate delineations of their field beyond deflationary ones such as “language evolution researchers are interested in the processes that led to a qualitative change from a non-linguistic state to a linguistic one” (Scott-Phillips & Kirby, 2010). One reason for this is probably that the discipline to be defined is itself a moving target, as it keeps evolving as new research foci emerge, and as basic theoretical assumptions keep changing. In the spirit of Jackendoff’s (2010) famous dictum that “Your theory of language evolution depends on your theory of language”, the scope of language evolution research will differ depending on what one thinks it is that evolved. As such, it can make sense to try answering such questions in a bottom-up manner, i.e. to move beyond intuitive stipulations and provide a data-driven characterisation of language evolution as a field of research. Previous

research (e.g. Bergmann & Dale, 2016) has already used quantitative data analysis to track the evolution of papers presented at the Evolang conference. We follow up on this line of research by providing a further ‘snapshot’ of the current state of the field, analyzing submission data from the Joint Conference on Language Evolution (JCoLE 2022, 5–8 September 2022, Kanazawa, Japan – see Section 3). As a one-off event combining different conference series and research communities, JCoLE very likely offers the best practically accessible representation of the current shape of the field of language evolution.

In the remainder of this paper, we will first give a brief overview of previous research (Section 2), and then move on to basic data on the conference (Section 3), and our case study (Section 4). Section 5 offers a discussion of the results and a conclusion.

2. Related research

Two existing bibliometric analyses of the Evolang conference have already shed light on the field of language evolution and its development since the mid-2000s. A pioneering study by Bergmann & Dale (2016) analysed the published proceedings of the Evolang conferences from 2006 to 2016, using Latent Dirichlet Allocation (LDA) to detect clusters of topics in the data. In addition, they used collaborations between authors to construct an authorship network and to identify interconnections between authors and topical clusters. Drawing on these data, they used centrality measures to detect the most influential nodes within the network. From this analysis, the cluster related to iterated learning and cultural evolution emerged as the “hub cluster” (Bergmann & Dale, 2016) for Evolang. In addition, they showed that “comparative studies with primates are an important interconnector between authors and communities, while the cluster covering linguistic approaches is interspersed and well represented throughout the conference.” (Bergmann & Dale, 2016). Similarly, a study of Evolang proceedings from the years 2004 to 2020 by Wacewicz et al. (2023) found that “greater betweenness centrality scores of primatologists and gesture researchers suggests those two research areas may be particularly important interdisciplinary connectors”. Using a topic modelling approach similar to that by Bergmann & Dale, they also show that the distribution of topics has undergone some subtle changes, including a vast diversification of topics and a growing importance of empirical methods.

Methodologically, the present study partly follows in the footsteps of Bergmann & Dale (2016) and Wacewicz et al. (2023). We combine a quantitative analysis of submission ratings with topic modelling; as we focus on one specific conference, we refrain from analyzing collaborations between authors, which

arguably makes more sense when taking a thoroughly longitudinal perspective. Importantly, our focus on JCoLE 2022 provides several benefits. While previous studies were limited to the *Evolang* conference series and to published proceedings, we take advantage (i) of the unique nature of the JCoLE 2022 conference as explained in Section 3 below, (ii) of direct measures of perceived relevance to the field, and (iii) of all submitted content, rather than only the published papers. Those differences in approach are due to our study having a different main question: What type of studies are typically considered by the scientific community to be within the scope of the field of language evolution?

3. The JCoLE 2022 conference

The Joint Conference on Language Evolution (JCoLE 2022; conference chairs: K. Okanoya, T. Hashimoto, & K. Fujita; proceedings: Ravnani et al., 2022), held 5–8 September 2022 in Kanazawa, Japan, was a conference of special status to the field of language evolution. It was a unique event jointly organised by all three main conference series in language evolution: *Evolang* (the main conference in the field), *Protolang* (a smaller conference, but well-established esp. in Europe), and the *Evolinguistics* project headed by Kazuo Okanoya, which brings together language evolution researchers from Japanese universities. After the 2020 edition of *Evolang* had been cancelled due to the Covid-19 pandemic, and the 2021 edition of the *Protolang* conference had taken place only virtually, JCoLE was the first in-person conference in the field in years. JCoLE 2022 was also the first conference on language evolution with both on-site and online participation options. It offered a sign language translation service (American Sign Language / International Sign) to make the conference accessible to a larger research community. It was organised outside of Europe, which previous research suggests may be conducive to a greater geographical diversity of submissions (Waciewicz et al., 2023). In the absence of a scientific society dedicated to studies in language evolution, researchers self-identifying with this field typically see the conference series mentioned above as a reference point; so in combining them, JCoLE 2022 arguably managed to bring together the language evolution community in all its breadth and diversity. For all the above reasons, it is reasonable to assume that the scope of submissions to JCoLE 2022 makes for a good representation of the scope of current language evolution research, affording a privileged source of insight into this field.

4. The study

What research counts as “language evolution research”? Which topics belong to the core of the research field, and which are rather seen as belonging to its periphery? To address these questions, we use aggregated, anonymized data of JCoLE 2022 submissions and the peer-review results, paying particular attention to the relevance scores that were obtained from the reviewers, as well as to the overall scores given by the reviewers.

4.1 Dataset

Submissions to the conference were invited in two formats – two-page abstracts or six- to eight-page full papers – each with a short summary (up to 150 words), author information, keywords, submission’s main research field (single choice from a list), and an optional supplementary document. The reviewing process was managed by the scientific committee of 12 researchers (5 female, 7 male) with different research backgrounds, who matched the submissions to reviewers selected from a pool of over 300 scholars of various research fields. Each submission was evaluated by 2 to 5 reviewers (mode: 3). The review process was double-blind (see Cuskley et al., 2020 for the implications of double-blind reviewing at Evolang). Each review consisted of a descriptive evaluation together with a numerical evaluation of “Overall quality” on a 7-point scale (from -3 to 3) and of “Topic relevance” on a 6-point scale (1 to 6).

Our dataset consists of 697 reviews by 187 reviewers on 255 submissions to JCoLE. 100 submissions were accepted for oral presentation, 122 for a poster presentation (including online posters), 33 submissions were rejected. Out of the 33 rejected papers, six were desk-rejected after careful discussion within the scientific committee for not meeting the scope of the conference (see Section 4.2.4), and were not sent to reviewers. The data were carefully anonymized, and clearance for the use of the data was obtained from the Data Protection Officer of Nicolaus Copernicus University in Toruń.

4.2 Analyses and results

4.2.1 *Fields of research*

We grouped all submissions by the main research field, as indicated by the submitting author. This first coarse subdivision gives an idea of the multidisciplinary present at the conference. The number of received submissions per field showcases the background of authors and, indirectly, which main fields consider

“language evolution” a pertinent research topic. The distribution of relevance scores might provide a general indication of how the language evolution community sees the current and future scope of the field.

The left panel in Figure 1 illustrates the proportion of submissions belonging to the individual topics, showing first of all considerable differences in the representation of the core research areas mentioned on the Evolang website – which are: anthropology, archeology, artificial life, biology, cognitive science, genetics, linguistics, modelling, paleontology, physiology, primatology, and psychology. Perhaps unsurprisingly, more than a third of all submissions (37.5%) fall into the linguistics category. Cognitive science is the second most chosen topic, namely in 16.9% of all submissions. These two fields together make up more than 50% of all submissions. Modelling, psychology, primatology, neuroscience, biology follow in this order. Each represents 5–10% of the submissions and together they make up almost 40% of all submissions.

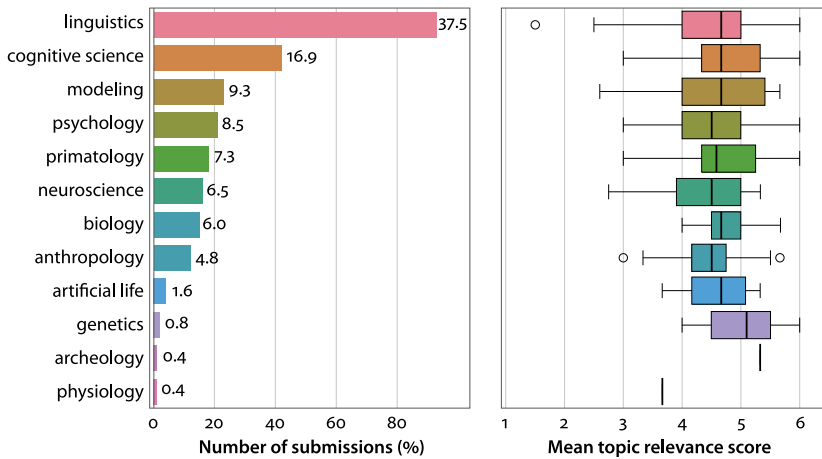


Figure 1. Absolute number of submissions (left) and reviewers’ relevance score (right) by main research field of the submission. The boxplot shows the median and the interquartile range for each category; as mentioned in Footnote 1, the data are based on mean relevance scores calculated for each paper

When it comes to the relevance scores given by reviewers, the data does not show very strong differences between fields. Figure 1 (right panel) shows the distribution of mean topic relevance scores per paper across the different main research fields.¹ Nonetheless some trends emerge: Focusing on research fields

1. We calculated the mean relevance score for each paper; the boxplots in Figure 1 and 3 are based on this summarized data, rather than the raw data for each individual review; this

with more than ten submissions, there is, as the box plots in Figure 1 (right panel) show, higher consensus in classifying research from biology as relevant; other fields such as neurosciences or modelling, by contrast, show a broad range of relevance ratings. As could be expected given the large number of submissions under the umbrellas of these two categories, this higher dispersion is especially true for the many submissions classified under linguistics and cognitive science, but also for e.g. primatology.

To check whether the topic relevance assessment differed more strongly for papers in some main research fields than in others, we also calculated the standard deviation between the topic relevance scores for each submission; Figure 2 shows the distributions of the standard deviations. Again, we do not see clear differences apart from a few outliers that are probably due to the low sample size in the respective categories, which indicates that the disagreements between the reviewers may be more due to the scope of individual papers than to perceived limits of the field itself.

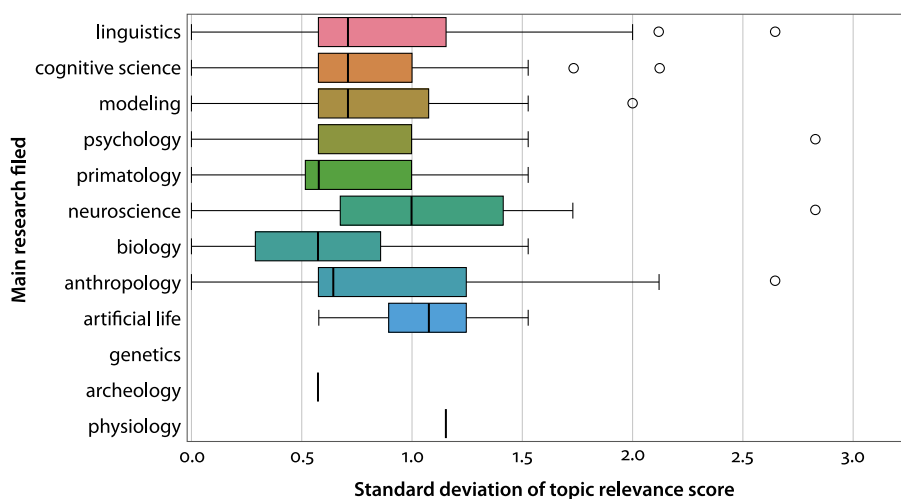


Figure 2. Standard deviations of the topic relevance scores. The standard deviation was calculated separately for each paper; the boxplots show the medians and the interquartile range

It can be debated of course how informative these indications of the main research field are, given that authors could only choose one field, and most submissions belong to at least two of the research fields. Also, some of the items that

ensures that papers with more reviews do not have a bigger impact on the displayed distributions than papers with fewer reviews.

can be chosen as the main research field pertain to scientific disciplines, while others pertain more to the methodological approach (e.g. *modelling*). As such, it would be interesting to compare these results with those from a more flexible typology if future editions of the relevant conference series decided to work with a revised set of main research fields with a multi-choice option.

4.2.2 *Topics of research*

Figure 3 shows the distribution of relevance scores for the keyphrases that authors could enter for their contributions. Only keyphrases used in 5 or more papers are displayed. In line with the results of previous research, which emphasised the centrality of cultural evolution research for current research on language evolution (see Section 5), *iterated learning* emerges as the keyphrase with the highest relevance score, followed by *word order* and *typology*. On the lower end of the scale, we find terms like *pragmatics*, *language change* and *language acquisition*. This shows that research on the dynamics of specific extant languages in modern humans seems to be perceived as less relevant than work that addresses the dynamics of language or language-like communication systems on a more abstract level, e.g. via communication games or by generalising over a large set of different languages. At the same time, it should be emphasized that the differences are fairly small and that hardly any paper has a mean relevance score below 3.

4.2.3 *Topic modelling and community detection*

Following the lead of previous scientometric work, we used topic modelling to group the submissions in a bottom-up way. We worked with 150-words summaries that the authors had entered for each submission; following a common procedure in topic modelling, we reduced inflected forms to their word stems using Porter stemming (Porter, 1980), as implemented in the NLTK Python library (Bird et al., 2009). We calculated their tf-idf score vectors (term frequency-inverse document frequency), a statistic that is commonly used to gauge how important a word is for a group of documents (Silge & Robinson, 2017). We calculated the pairwise cosine similarity between documents, and used these similarity measures as the edge weights (i.e., the strength of a network link) in an undirected graph. On this graph, we ran an off-the-shelf community detection algorithm (Clauset-Newman-Moore greedy modularity maximisation, implemented in the NetworkX Python library; Hagberg et al., 2008) to find 15 disjunct groups of submissions,² clustered by the terms used in these author-provided summaries.

2. The number of groups is set by the analyst, depending on the desired coarseness of the analysis. Here, setting the number to 15 groups strikes a decent balance between detail and interpretability.

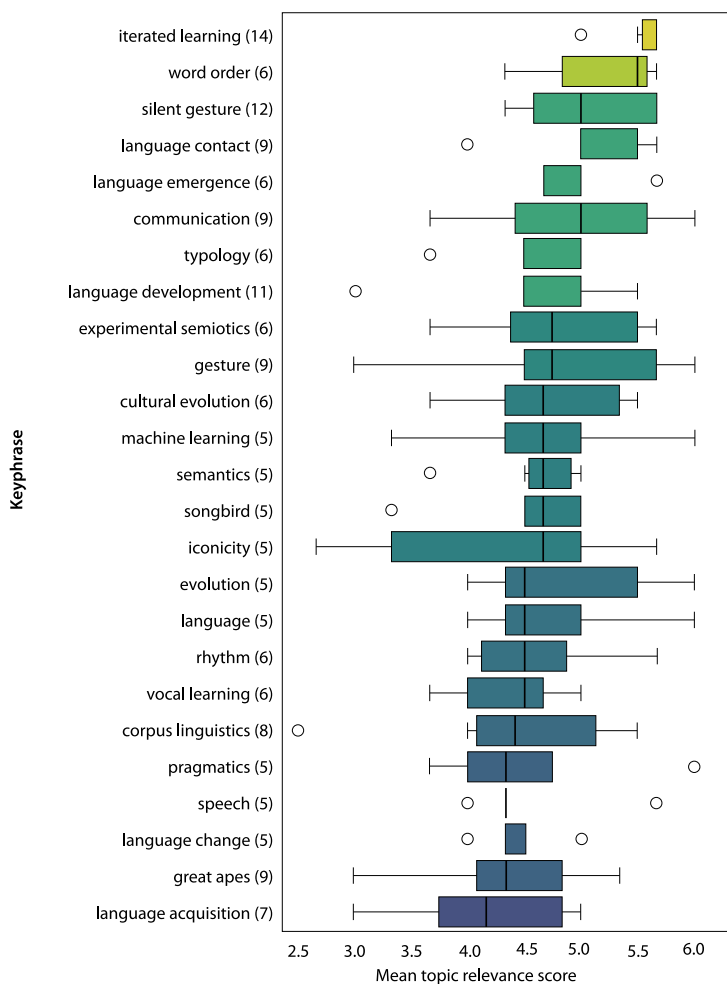


Figure 3. Relevance score by keyphrase, in descending order of the median. Only keyphrases used 5 or more times are taken into account. The numbers in brackets indicate how many different papers used the keyphrase. As in Figure 1, the data are based on the mean relevance score calculated for each paper

To get an idea of the topics of submissions within each group, we removed stopwords (i.e. words that are very frequent in English anyway and as such do not contribute significant information) from the summaries and again used the tf-idf to determine the most relevant words in each group of submissions. Figure 4 shows the detected groups with the 10 most relevant terms, specifying their mean relevance scores and the composition of the “main research field” of the group’s submissions (see Section 4.2.1).

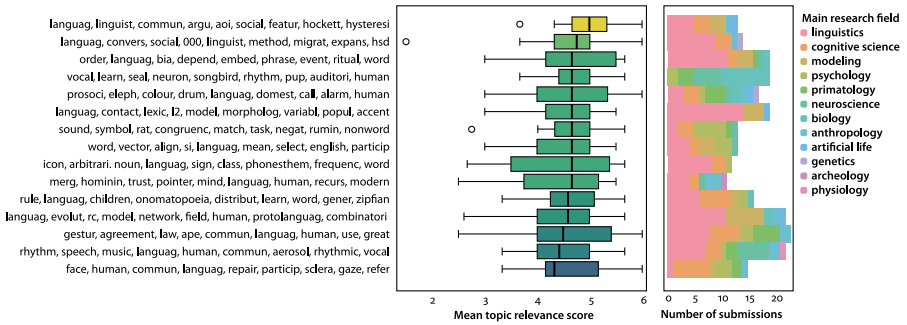


Figure 4. Groups of submissions detected based on the 10 most relevant terms. The boxplots (left) show the distribution of mean relevance scores within each group, and the barplots (right) show the composition of the “main research field” of the group’s submissions (see Section 4.2.1)

In the results we can see a mix of traditional linguistic topics (e.g. morphology, language contact), as well as relatively newer topics (sign languages, gestures), or ones that are only beginning to capture the attention of linguists (rhythm and music, vocal learning). This pattern of results suggests that the interdisciplinarity of language evolution has the potential to inspire and invigorate linguistics. A main observation is that, once again, differences between groups in the assessed relevance of the submissions are small, testifying the inherent multidisciplinary nature of language evolution as a research field. The fact that no research topic is entirely dominated by submissions from a single main research field further highlights this point.

4.2.4 Further analyses

In this subsection, we take a closer qualitative look at the submissions that received the highest and the lowest relevance scores, respectively. It should be noted that the overall distribution of topic relevance scores was skewed to the higher end of the scale. This is also the case for the overall evaluation scores, where the proportion of negative scores is, however, higher overall (see Figure 5). This indicates that most submissions were identified as topical, and only very few submissions were deemed not relevant for the conference topic at all.

Submissions with the highest and lowest relevance scores

Still there were only six papers that unanimously received the highest score (6) from all three reviewers. Their topics included gestural communication, genetic bases of language, iterated learning, and the evolution of non-lexical units. The topics of iconicity, iterated learning experiments, constituent order in spoken and signed languages, the evolution of grammatical structure on the one hand and

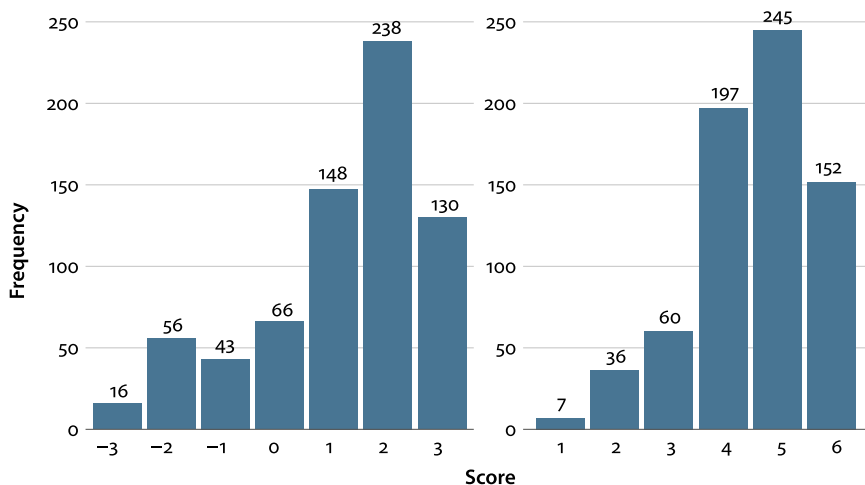


Figure 5. Distribution of overall evaluation scores (left) and topic relevance scores (right). Each datapoint represents one review

meaning on the other in natural language, as well as studies on cross-species comparison feature prominently among the papers that have received the highest score from at least two reviewers. This pattern of results again illustrates the inherent disciplinary breadth of the field.

Turning to the submissions with the *lowest* relevance scores, 36 submissions received at least one relevance score of 2 or 1, i.e. were deemed minimally relevant by at least one reviewer. Only six papers, however, received such low scores by more than one reviewer. Most of those papers deal with statistical or distributional regularities in today's languages (sometimes limited to individual languages).

Rejections by the scientific committee

Finally, several papers were desk-rejected for not meeting the formal and scientific quality or the scope of the conference. These papers were rejected for one of the following reasons: (1) papers not meeting formal and scientific criteria (e.g., missing reference list), (2) papers presenting purely linguistic analysis of modern language(s) without clearly stating its relevance to language evolution; and (3) papers without adequate empirical support. Topic-wise, those papers can be categorized as follows: (1) language variation, language contact, and language change; (2) specific features of (proto-)language(s); and (3) postulating direct links between linguistic features and certain biological traits.

Topic relevance and overall evaluation

As could be expected, there is a high correlation between topic relevance and overall evaluation in our data (Kendall's $\tau = 0.48$, $z = 15.56$, $p < 0.001$). However, Figure 5 illustrates an interesting pattern of more papers receiving an overall negative evaluation despite being identified as topical (bottom-right corner) than a positive evaluation despite perceived limited relevance (upper left corner). At first glance, we could expect most of the relevant but low-scoring submissions to propose speculative scenarios of language origins, something that the field used to be infamous for (cf. e.g. Fitch, 2017). However, a qualitative look at the papers that received low overall scores but high relevance ratings reveals that with a few exceptions, most of them do not offer 'just so stories', but rather describe coherent theoretical models that are consistent with existing hypotheses about language origins or describe empirical work, but in a way that did not convince the reviewers.

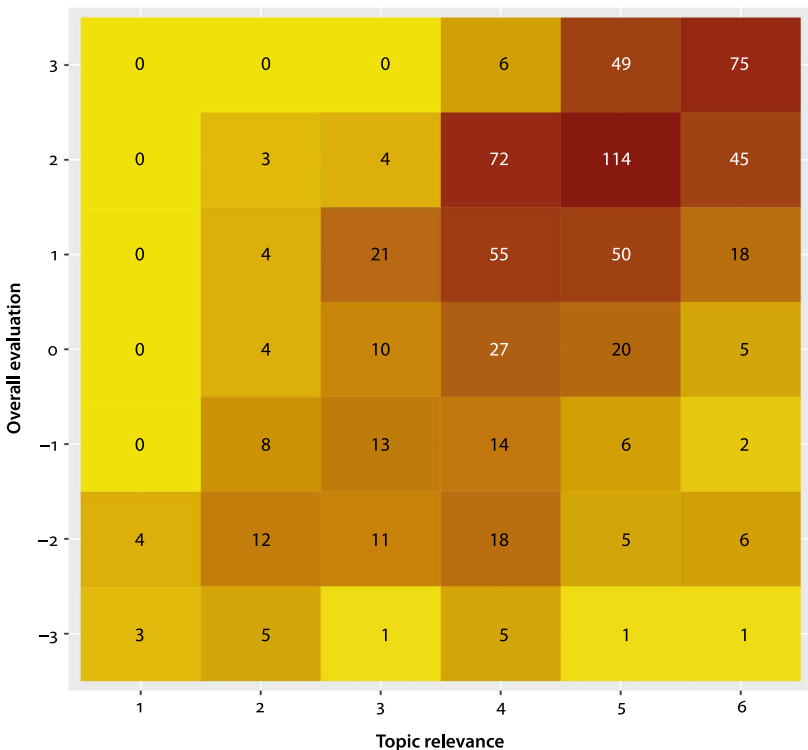


Figure 6. Distribution of topic relevance and overall evaluation scores. (Each datapoint is one review; the darker the color, the more data points)

Empirical vs. theoretical submissions

Finally, we looked into the distribution of theoretical versus empirical submissions, and how they fared in the review process. All submissions were coded by two independent raters for whether they are empirical (i.e., based on authentic or elicited data), modelling-based, or predominantly theoretical. Inter-rater agreement was high, with Cohen's $\kappa=0.7$, which qualifies as "substantial agreement" according to Landis & Koch (1977). In the following analysis, we take into account only those data points for which both raters chose the same annotation (empirical $n=134$, modelling $n=20$, theoretical $n=58$). As the left panel of Figure 7 shows, empirical papers were much more likely to be accepted (especially for podium presentations) than all other types, while theoretical papers were more likely to be rejected (or to be accepted as an online poster; note that onsite posters were offered to submissions with a higher evaluation score, while online posters were offered to submissions with a lower evaluation score). As for topic relevance, there is not a large difference between the types; although the theoretical contributions were evaluated slightly lower than the empirical and esp. modelling ones (Figure 7, right panel), the difference was negligible (mean topic relevance score for theoretical papers: 4.46; $sd=1.16$; for modelling papers: 4.61, $sd=1.22$; for empirical papers: 4.61, $sd=1.09$).

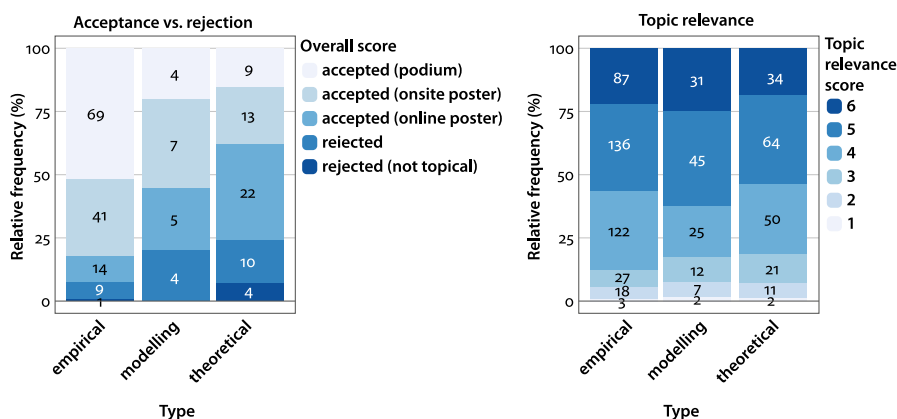


Figure 7. Paper acceptance rates (on the left) and distribution of topic relevance scores (on the right) for empirical, modelling, and theoretical papers. Counts refer to papers on the left panel, and to reviews on the right panel. Only papers which both annotators classified in the same way were taken into account

5. Discussion and conclusion

The results of our study allow for valuable insights regarding the question of what the language evolution research community regard as being part of their research field. They align with and substantiate the main results of previous studies such as Bergmann & Dale (2016) or Waciewicz et al. (2023): we see a broad diversity of research topics, with some being better-represented in terms of numbers, but only very weak differences in terms of the reviewers' relevance assessments. In terms of pure quantity of submissions, our data show a certain bias towards the investigation of cultural language evolution as well as studies comparing communicative signals in non-human animals to human linguistic communication. These results strongly align with those obtained by Bergmann & Dale (2016). The results also indicate that linguistic approaches are deemed particularly relevant for language evolution when taking a broad-scope and general perspective, whereas studies confined to individual languages are considered less relevant, as the manual inspection of papers that received the lowest relevance scores has confirmed. Several areas that have been traditionally perceived as key for language evolution research (see e.g. the contributions to Gibson & Tallerman, 2011) – including e.g. archaeology, palaeoanthropology, and genetics – are underrepresented, which may to some extent result from the resource-intensiveness of studies in those fields. Cognitive science was the research field with the second highest relevance score; this may reflect the interdisciplinarity of language evolution research, given that cognitive science may be taken to comprise several individually relevant disciplines, such as linguistics, philosophy, psychology, neuroscience, anthropology, and computer science (Miller, 2003).

Our study has also shown that most submitted papers received very high relevance scores and the difference between topic groups in the assessed relevance scores is relatively small. This can likely be accounted for by a combination of multiple explanations. First of all, it might be indicative of the breadth of the field. For example, the results of our topic analysis in 4.2.2 and 4.2.3 indicate that “language evolution” is focused not only on the human level, but also comprises other communication systems, including non-human animal acoustic and gestural communication. It could also indicate that the authors submitting to a conference on language evolution research tend to have a clear impression of the scope of the field. Through the established conference series such as *Evolang* and *Protolang* that have already provided a platform for interdisciplinary discussions for a long time, researchers may have developed some implicit common understanding of what counts as language evolution research.

Finally, our investigation of the evaluation of study types (i.e., empirical, modelling, and theoretical work) shows that the community strongly values empiri-





cal work. This indicates that language evolution is considered as an empirically tractable phenomenon by the community members. There is also a substantial and growing tendency to use existing data, with roughly 25% of submissions reporting studies working with data already available from databases and corpora (rather than collecting new primary data). As for experimental work, empirical research using communication-game setups is considered highly relevant despite the theoretical concerns with using modern humans as participants.

Our results offer a simple snapshot of a dynamically developing field. The faced-paced nature of progress in language evolution research makes it especially interesting and inviting to conduct similar-type analyses again in the future – taking advantage of considering all submissions and their relevance scores, rather than just accepted work – to enable longitudinal comparisons and the identification of trends over the years.

Funding

SW was supported by the Polish National Science Center under grant agreement UMO-2019/34/E/HS2/00248. EDR was supported by the Portuguese national funding agency for science, research, and technology (SFRH/BD/138406/2018). RA saw supported by the projects Mext 4093 (Evolinguistics), JP22Ho4897 and JP17Ho6378-383.

References

- Bergmann, T. & Dale, R. (2016). A Scientometric Analysis Of Evolang: Intersections And Authorships. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér & T. Verhoef (Eds.), *The Evolution of Language* (pp. 79–87). Proceedings of the 11th International Conference.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python: analyzing text with the natural language toolkit*. O’ Reilly Media, Inc.
-  Cuskley, C., Roberts, S. G., Politzer-Ahles, S., & Verhoef, T. (2020). Double-blind reviewing and gender biases at EvoLang conferences: An update. *Journal of Language Evolution*, 5(1), 92–99.
-  Dediu, D., & de Boer, B. (2016). Language evolution needs its own journal. *Journal of Language Evolution*, 1(1), 1–6.
-  Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review* 24(1). 3–33.
-  Gibson, K. R., & Tallerman, M. (Eds.). (2011). *The Oxford Handbook of Language Evolution*. Oxford University Press.
- Gong, Y., Chen, X., & Zhang, M. (2014). Evolutionary linguistics in the past two decades: The 10th international conference on language evolution. *Journal of Chinese Linguistics*, 42, 499–530.

- doi** Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161), 581–598.
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In Varoquaux, G., Vaught, T., & Millman, J. (Eds.), (pp. 11–15). Proceedings of the 7th Python in Science Conference (SciPy2008)
- doi** Hartmann, S. (2020). Language change and language evolution: Cousins, siblings, twins? *Glottology*, 11(1), 15–39.
- Haspelmath, M. (2016). The evolution (or diachrony) of “language evolution” [Billet]. *Diversity Linguistics Comment*. <https://dlc.hypotheses.org/894>
- doi** Jackendoff, R. (2010). Your theory of language evolution depends on your theory of language. In Larson, R. K., Déprez, V. & Yamakido, H. (eds.), *The evolution of human language: Biolinguistic perspectives*, 63–72. Cambridge: Cambridge University Press.
- doi** Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174.
- Lewontin, R. C. (1998). The evolution of cognition: Questions we will never answer. In *Methods, models, and conceptual issues: An invitation to cognitive science, Vol. 4*. (pp. 106–132). The MIT Press.
- doi** Mendívil-Giró, J. L. (2019). Did language evolve through language change? On language change, language evolution and grammaticalization theory. *Glossa: A journal of general linguistics*, 4(1), 124.
- doi** Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, 7, 141–144.
- doi** Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.
- doi** Ravignani, A., Asano, R., Valente, D., Ferretti, F., Hartmann, S., Hayashi, M., Jadoul, Y., Martins, M., Oseki, Y., Rodrigues, E. D., Vasileva, O., & Waciewicz, S. (Eds.). (2022). *The evolution of language: Proceedings of the Joint Conference on Language Evolution (JCoLE)*. The Joint Conference on Language Evolution (JCoLE).
- doi** Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14(9), 411–417.
- Silge, J. & Robinson, D. (2017). *Text mining with R. A Tidy approach*. O’Reilly.
- doi** Waciewicz, S., Żywicznyński, P., Hartmann, S., Pleyer, M., & Benítez-Burraco, A. (2020). Language in Language Evolution Research: *Biolinguistics*, 14, 59–101.
- doi** Waciewicz, S., Sibierska, M., Placiński, M., Szczepańska, A., Poniewierska, A., Ng, Y. Y., & Żywicznyński, P. (2023). The scientometric landscape of Evolang: A comprehensive database of the Evolang conference. *Journal of Language Evolution*.
- doi** Żywicznyński, P. (2018). *Language Origins: From Mythology to Science*. Peter Lang.

Address for correspondence

Stefan Hartmann
German Department
Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1
40225 Düsseldorf, NRW
Germany
hartmast@hhu.de

Biographical notes

Stefan Hartmann. Professor at Heinrich Heine University Düsseldorf. He investigates language dynamics: variation, change, acquisition, evolution.

Sławomir Wacewicz. Associate Professor at the Nicolaus Copernicus University in Toruń. Interested in Evolutionary Psychology and Cognitive Evolution, his research focuses on language evolution and evolution of the mind.
wacewicz@umk.pl

Andrea Ravignani. Professor at the Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy. Formerly leader of the Comparative Bioacoustics Group at the Max Planck Institute for Psycholinguistics. He investigates speech, rhythm, and communication in mammals, especially seals and humans.
andrea.ravignani@uniroma1.it

Daria Valente. Researcher at the University of Torino, investigating vocal communication in primates, with a particular focus on lemurs.
daria.valente@unito.it

Evelina Daniela Rodrigues. PhD Student at ISPA–Instituto Universitário. Her research explores the development of communication and accommodation in human and chimpanzee gestures.
erodrigues@ispa.pt

Rie Asano. Works at the Advanced Comprehensive Research Organization, Teikyo University, Tokyo, Japan, and previously was researcher of Musicology of the University of Cologne. She studies the relationship between linguistic syntactic processing and musical rhythm, computational neurocognitive modeling, and computational evolutionary neuroscience.
rie.asano@uni-koeln.de

Yannick Jadoul Researcher at the Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy, focusing on computational modeling and artificial intelligence to study rhythm, speech, and animal vocalisations.

Yannick.Jadoul@uniroma1.it

Publication history

Date received: 22 April 2023

Date accepted: 26 February 2024