GLOSSA
psycholinguistics

eScholarship
University of California

A decade of language processing research: Which place for linguistic diversity?

Aymeric Collart, Institute of Linguistics, Academia Sinica, Taiwan, aymeric.collart@gmail.com

This paper surveys the linguistic diversity in psycholinguistic and neurolinguistic research by examining the languages under investigation in major international conferences from 2012 to 2023. The results showed that these studies are highly skewed towards English in particular and Indo-European languages in general. However, the overall number of languages, as well as the number and proportion of Indo-European (other than English) and non-Indo-European languages increased over time, indicating that language processing research is becoming more and more diversified. This typological bias was also found in the inspection of specific linguistic phenomena: (a) morphosyntactic alignment, richness of case morphology, canonical word order, and (b) temporal concepts. The analyses of typological bias at the general and specific levels indicate that there are gaps in various topics, and these can be filled by including more non-Indo-European languages in the investigation process. In addition, a sociolinguistic bias in language processing research emerges as the languages investigated are more often 'Western' languages with more than one million speakers and a shared written form. These results reflect the numerous challenges encountered when conducting experiments on less familiar languages, such as the geographic difficulty finding participants speaking these languages, the need of institutional support, as well as the difficulties in setting up collaborations with native speakers, among others.

1. Introduction

Linguists largely agree that languages are a product of the mind, regardless of which one we speak or sign. Despite this claim, it is quite evident how different languages can be at every level – phonetic, phonological, morphological, lexical, semantic and syntactic. Being a product of the mind, the nature of the cognitive mechanisms underlying language production and comprehension has been the focus of several scientific disciplines, such as psycholinguistics and neurolinguistics. However, researchers have expressed concerns that language processing research is facing challenges from language diversity (Blasi et al., 2022; Bornkessel-Schlesewsky & Schlesewsky, 2016). Can we assume general processing mechanisms which could apply to all languages? Do language-specific grammatical characteristics lead to different processing patterns, and if so, to what extent?

The present study evaluates to what extent the field can address these overarching questions by surveying the diversity of language processing research in terms of linguistic typology through the examination of the abstracts of studies presented in seven conferences between 2012 and 2023. In other words, to what extent is language processing research diverse in terms of the inclusion of different languages, and can we observe any change over time? This paper is organized as follows. First, the question of the need for linguistic diversity in language processing research is discussed in Section 2. Section 3 introduces previous quantitative studies conducted on this issue, and the present survey (including the methodology and the results) is reported in Section 4. A discussion of the results is given in Section 5, especially concerning the factors challenging the inclusion of less common languages in language processing research. Section 6 concludes the paper.

2. Does language processing need linguistic diversity?

Before investigating how diverse language processing research is, it is crucial to consider the question of whether this domain needs linguistic diversity.[1] Indeed, if we assume that languages are the product of the mind, we might argue that the same cognitive operations are at stake when it comes to language production and comprehension, regardless of the language, and linguistic differences are just different realizations of these shared operations. The inspection of influential neurolinguistic models of sentence processing may point in this direction. The Memory, Unification, Control model proposed by Hagoort (2003, 2016) includes languages which are exclusively Germanic (Indo-European): English, German and Dutch. It must be noted that Hagoort (2003, 2016) does not make explicit reference to the importance of including languages from different language families. Similarly, Friederici's (2011, 2016) model includes observations from six languages (English, German, Dutch, Hebrew, Japanese and Thai), covering four language families. This number of languages seems quite small when compared to the

---

[1] See also Jaeger and Norcliffe (2009) and Norcliffe et al. (2015) for a discussion of this issue.

estimated 7000 languages of the world (Evans & Levinson, 2009; Kidd & Garcia, 2022), and questions may be raised regarding the extent to which these models can indeed be generalized to all languages. Friederici (2011, p. 1386) does address this concern, arguing that relying on a rather small sample of languages that are representative of different families does not prevent one from generalizing the results, as seems to be the case concerning syntactic structure building, which activates similar processing patterns across the six languages mentioned above. At the same time, Friederici (2011, p. 1386) does not refute the fact that for some specific topics, such as the processing of argument structure, inclusion of a larger variety of languages may be needed.

Does language processing research need observations from various languages to be valid? The answer to this question needs the same type of empirical evidence, i.e., psycholinguistic and neurolinguistic experiments must be conducted on a large pool of languages in order to determine to what extent processing patterns are stable or not across different types of languages. As Malik-Moraleda et al. (2022, p. 1018) put it, "probing human language in all its diverse manifestations is critical for […] understanding the cognitive and neural basis of different solutions to communication demands", as well as "characterizing the processing of unique/rare linguistic properties and fostering diversity and inclusion in language sciences". Based on these observations, Malik-Moraleda et al. (2022) examined which brain areas were activated during general sentence comprehension, using the functional magnetic resonance imaging (fMRI) technique, in which the participants had to read sentences or listen to passages of stories (the baseline conditions involved unintelligible sentences). The particularity of this study is that Malik-Moraleda et al. (2022) conducted the same experimental protocol in 45 languages, covering twelve language families.[2] The results showed that despite cross-linguistic differences, similar brain areas and processing patterns were found across these languages when it comes to general comprehension mechanisms. But the investigation of other topics may reveal that cross-linguistic differences lead to qualitatively different processing patterns. This is the case, for example, for the comprehension of argument structure, briefly discussed below.

Some models of sentence comprehension do systematically investigate the processing of cross-linguistic variation. For instance, the Competition Model, first proposed by E. Bates et al. (1982) in their study comparing the factors influencing word order processing in English and Italian, now provides generalizations based on a sample of 18 languages (MacWhinney, 2022). Another example is the extended Argument Dependency Model, which seeks neurotypological validation by making processing generalizations based on different languages belonging to typologically unrelated families (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky,

---

[2] Nevertheless, it must be noted that Indo-European languages (31 out of 45) clearly outnumbered other language families, and also that eight language families had just one representative: Austro-Asiatic, Austronesian, Japonic, Koreanic, Atlantic-Congo, Sino-Tibetan, Turkic and Isolate (Malik-Moraleda et al., 2022, p. 1015).

2013, 2016). As an example of neurocognitive variation investigated within this model, Bornkessel-Schlesewsky et al. (2011) carefully selected six languages from four language families based on grammatical characteristics influencing how the argument structure of a predicate is built, summarized in **Table 1** (adapted from Bornkessel-Schlesewsky et al. (2011, p. 146)).[3]

Table 1: Languages under investigation in Bornkessel-Schlesewsky et al. (2011, p. 146).

Language family	Language	Case morphology	Basic word order	Cue ranking
Indo-European (Germanic)	English	Poor	SVO	WO > animacy
	Dutch	Poor	SVO/SOV	WO > animacy
	German	Rich	SVO/SOV	animacy > WO
	Icelandic	Rich	SVO	?
Turkic	Turkish	Rich	SOV	animacy > WO
Sino-Tibetan	Mandarin Chinese	Poor	SVO	animacy > WO

They used the event-related potentials technique (ERP) to investigate the phenomenon of *semantic reversal anomaly*, as exemplified in (1), which was assumed to be sensitive to the grammatical characteristics summarized in **Table 1**.

(1)   a.   *No semantic reversal anomaly*
          The hungry boy was devouring the cookies. (Kim & Osterhout, 2005, p. 208)
    b.   *Semantic reversal anomaly*
          The hearty meal was devouring the kids. (Kim & Osterhout, 2005, p. 208)

In (1a), the subject of the verb is also the agent of the predicate, based on the word order of the sentence, which is a major clue in English to determining the argument structure of the sentence. Based on one's world knowledge, it is ordinary for hungry boys to devour cookies, and the sentence is acceptable. In (1b), the subject and object nouns have been reversed. Even if the sentence is syntactically correct, it violates one's world knowledge, and crucially, it could be perfectly acceptable if the subject and object nouns were reversed (i.e., if their semantic roles were exchanged). The experimental results from these six languages, taken together, revealed that different processing patterns were at stake depending on grammatical characteristics concerning the attribution of semantic roles, indicated by the elicitation of a N400, P600 or biphasic N400-P600 response.

---

[3] The English and German data come from previous studies (English: Kim & Osterhout, 2005; German: Schlesewsky & Bornkessel-Schlesewsky, 2009).

The inclusion of understudied languages in language processing research can indeed provide new insights, since such languages may exhibit peculiar grammatical characteristics. This is exemplified here with Formosan languages, the Austronesian languages spoken in Taiwan (Collart, 2024). In the case of the processing of argument structure, Formosan languages differ from the languages investigated by Bornkessel-Schlesewsky et al. (2011) in that they exhibit a morphological voice system not found in Indo-European, Sino-Tibetan or Turkic languages. The predicates (generally found in the sentence-initial position) are inflected with voice morphemes, such as actor voice or patient voice morphemes, which indicate the thematic role of the subject noun marked with nominative case. For instance, the noun referred to as the subject in the sentence is interpreted as the actor when the predicate is inflected with the actor voice morpheme, and as the patient when the predicate is inflected with the patient voice morpheme.[4] Yano et al. (2019), Ono et al. (2020) and Sato et al. (2020) took advantage of these characteristics in addition to the flexible word order in Truku Seediq (despite its canonical VOS word order) to investigate the interplay of two factors during the computation of the argument structure of a sentence: *conceptual accessibility* ('actor' role preceding 'patient' role is more easily processed) and *syntactic complexity* (canonical word order is easier to process than derived word order). They found that even if the factor of syntactic complexity is prominent, conceptual accessibility may also play a role under certain circumstances. Crucially for the present discussion, these two factors could be more easily teased apart thanks to the grammatical characteristics of Formosan languages, suggesting that language processing research can benefit from language diversity.

This section has highlighted the limitation of linguistic diversity in existing theory (with a focus on neurolinguistic models) and also reviewed some specific examples of the value of linguistic diversity in advancing theory. The next section turns to a more systematic effort to quantify linguistic diversity in the language processing literature.

3. Quantifying linguistic diversity in language processing

Researchers have highlighted a perceived lack of typological diversity in language processing research (Hawkings, 2007, p. 104). Making such a claim requires quantifying this amount of diversity, and two studies focused on this issue: Anand et al. (2011) and Kidd and Garcia (2022).

3.1 Linguistic diversity in psycholinguistic and neurolinguistic abstracts until 2011

Anand et al. (2011), after surveying the languages under study in 4550 abstracts of conferences and journals focusing on language processing and corpus exploration, found that even if a total

---

[4] This is a very simplified view of the voice system in Formosan (and more generally in Austronesian) languages. Other voice morphemes can be found, such as instrumental or locative voice morphemes, and the use of some voice morphemes may also interact with other factors, such as the aspectual class of the predicates (i.e., stative or non-stative predicates).

57 languages had been investigated, 85% of the studies were represented by just ten languages. These are given in **Table 2** (only the data of language experiments are reported).

Table 2: The ten most represented languages in language processing experiments (Anand et al., 2011).

Rank of representativity	Language	Family
1	English	Indo-European
2	German	Indo-European
3	Japanese	Japonic
4	French	Indo-European
5	Dutch	Indo-European
6	Spanish	Indo-European
7	Mandarin	Sino-Tibetan
8	Korean	Koreanic
9	Finnish	Finno-Ugric
10	Italian	Indo-European

These languages cover four families, but it is obvious that Indo-European languages are more represented than the others, as six languages out of ten belong to this family, as opposed to only one for the three other families. Anand et al. (2011) explain this underrepresentation of language diversity to the fact that psycholinguistic and neurolinguistic experiments may be challenging to conduct on understudied languages, which is a point more thoroughly discussed in 5.3.

3.2 Linguistic diversity in language acquisition experiments between 1974 and 2020

Kidd and Garcia (2022) conducted a similar survey, but focused on language acquisition studies published in articles in four journals ranging from 1974 to 2020. The 3310 studies taken into consideration involve 103 languages.[5] However, as the summary in **Table 3** of the ten most represented languages in this survey suggests, despite the greater number of languages than in Anand et al. (2011), these studies are largely biased: the ten most represented languages account for 84.41% of the articles, with most of them belonging to the Indo-European family. The same conclusion as in Anand et al. (2011) can be reached: language acquisition research is heavily biased towards some languages, and particularly languages which are typologically similar.

---

[5] Kidd and Garcia (2022, p. 707) indicate that they included 2826 articles in the analysis, but the frequency distribution of articles given in the appendix encompasses 3310 data points. The data in their appendix are reported here.

Table 3: The ten most represented languages in language acquisition studies (Kidd & Garcia, 2022).

Rank	Language	Family	Number of studies	Percentage
1	English	Indo-European	1790	54.08
2	French	Indo-European	215	6.50
3	Spanish	Indo-European	153	4.62
4	German	Indo-European	136	4.11
5	Dutch	Indo-European	115	3.47
6	Italian	Indo-European	114	3.44
7	Hebrew	Semitic	90	2.72
8	Mandarin	Sino-Tibetan	83	2.51
9	Japanese	Japonic	63	1.90
10	Russian	Indo-European	35	1.06

Nevertheless, changes between 1974 and 2020 must be noted. While the number of language acquisition articles grew over time regardless of the language under investigation, the pace was faster for non-Indo-European languages, and even faster for Indo-European languages other than English, when compared with the pace of studies published focusing on English, especially in the last twenty years (Kidd & Garcia, 2022, p. 714). In other words, even if the lack of linguistic diversity is obvious in language acquisition research, this pattern is changing but at a very slow pace.

3.3 A sociolinguistic bias in addition to the typological bias?

The surveys conducted by Anand et al. (2011) and Kidd and Garcia (2022) suggest that the main problem with the lack of typological diversity in language processing research is that relying on languages which belong to the same genetic family implies that these languages share similar characteristics. To some extent, this corresponds to what Dahl (1990) and Haspelmath (2001) refer to as "Standard Average European". Such languages have also been identified by some scholars as 'WEIRD', an acronym standing for languages spoken in 'Western, Educated, Industrialized, Rich and Democratic' societies (Heinrich et al., 2010). Because general theoretical assumptions are based on these languages, they have even been seen as "misleading" cognitive science and linguistics, understood here as having biased potential generalizations (Henrich et al., 2010; Majid & Levinson, 2010).

Kidd and Garcia's (2022) study shows that even if there is still a great gap between the number of English, Indo-European, and non-Indo-European studies, this gap has become smaller over time. The same observations have been made for theoretical linguistics. However, as Dahl (2015, p. 4)

puts it, "even after the Eurocentric bias has started to lose its grip on the choice of languages to be studied, there remains a bias that can be summed up with the acronym 'LOL'", where the first L stands for 'Literate' (languages with a writing system used by most of the speakers), O for 'Official' (whether it is the official language of a territory), and the second L for 'Lots of users' (languages with at least one million speakers). In total, 57 languages exhibit these features.

Can we also observe the same sociolinguistic bias in language processing research? A reanalysis of the data from Kidd and Garcia (2022) shows that among the 103 languages under investigation, 61 are characterized as non-LOL, and 42 as LOL. However, the 61 non-LOL languages covered 190 studies, or 5.74% of the corpus collected by Kidd and Garcia (2022), while the 42 other languages correspond to 94.26% of the studies. Therefore, there is not only a typological bias in language processing research, but also a second one which is sociolinguistic.

4. A decade of linguistic diversity in language processing conferences: Survey

4.1 Rationale for the present study

The surveys conducted by Anand et al. (2011) and Kidd and Garcia (2022) provided quantitative data on a phenomenon which has been widely perceived in experimental linguistics. However, these two studies focused on different domains: psycholinguistics for Anand et al. (2011) and child acquisition research for Kidd and Garcia (2022). The data provided by Kidd and Garcia (2022) allow a finer-grained view of the evolution over time of linguistic diversity in language acquisition research, while the data from Anand et al. (2011) take the form of a 'snapshot', which is not informative regarding changes over time. In addition, it is not clear whether the findings for language acquisition research hold for other psycholinguistic and neurolinguistic experiments. Indeed, the profiles of the participants differ, and acquiring data from children may introduce an additional challenge which may prevent researchers from conducting experiments in understudied languages.

Therefore, the present study surveys the diversity of languages in psycholinguistic and neurolinguistic experiments which are not restricted to language acquisition studies. Taking the data from Anand et al. (2011) as a starting point, this survey covers each year from 2012 to 2023 in order to track changes over time. Unlike Kidd and Garcia (2022), this survey investigates abstracts from different types of conferences and not journal articles. Indeed, experiments whose data are collected may not necessarily be published in journal articles, unlike conference abstracts, which may reflect studies at more preliminary stages, and this may thus affect the results of the survey.

4.2 Methodology

4.2.1 Corpora and annotations

Seven conferences in language processing research were selected. They were divided into three groups. The first group includes *major* conferences (in terms of number of studies presented),

namely (a) the CUNY Conference on Sentence Processing, renamed Annual Conference on Human Sentence Processing (HSP) since 2021, (b) the Architectures and Mechanisms for Language Processing conference (AMLaP), and (c) the Society for the Neurobiology of Language Annual Meeting (SNL). These three conferences, held annually and generally at different times of the year, host several hundreds of presentations across a broad range of methods. Therefore, they are particularly suitable for the present survey, as they provide a large amount of data. Abstracts from 2012 to 2023 were included in the survey. As previously mentioned, the starting year was selected based on the previous similar survey conducted by Anand et al. (2011).[6] The second group includes smaller conferences held in Asia, including (a) the Asian venue of AMLaP (two iterations in 2018 and 2023), (b) the International Conference on Theoretical East Asian Psycholinguistics (ICTEAP, four instances in 2017, 2019, 2021 and 2023), and (c) the South Asian Forum on the Acquisition and Processing of Language (SAFAL, held annually since 2020). The third group consists of one conference explicitly focusing on cross-linguistic research in language processing, the Crosslinguistic Perspectives on Processing and Learning conference (X-PPL, four instances in 2019, 2021, 2022 and 2023). This comparison will be helpful in understanding the effect of the geographical location as well as the main focus of the conferences on the diversity of the studies.

Data annotation was manually performed by looking through the abstract of each study. At first, the dataset only included annotations about (a) the conference name, (b) the year of the conference, (c) the presentation type (oral or poster presentation), and (d) the title of the study. The language(s) under investigation were retrieved by reading through each abstract manually. For example, if the study explored the issue of negative polarity processing in German, 'German' was annotated as the language under investigation. Dialectal variations were not considered, i.e., British English and American English were both annotated as 'English'. Sign languages were annotated as such, i.e., 'American Sign Language', 'British Sign Language', 'Italian Sign Language'. Studies involving several languages were entered as several lines in the spreadsheet (as many lines as the number of languages under investigation). Studies which focused on statistical methods, corpus exploration, or models at a more abstract level were annotated as 'NA', and studies using artificial language were annotated as such, too. Complications appeared concerning cross-linguistic studies and language acquisition/learning studies. Indeed, some abstracts referred to dozens of languages, such that it was not completely clear whether the data from these languages were collected for that particular study or whether they came from previous studies (these corresponded to the exclusion of only 28 datapoints). As for language acquisition/learning studies, only the language tested was taken into consideration. For example, if two groups of participants were recruited, i.e., native speakers of English and Japanese native speakers with English as their L2, and the experiment only included English stimuli, this study was annotated as investigating

---

[6] Only one booklet of abstracts was missing in this range (AMLaP 2013). Another reason to start from 2012 is a technical one, as previous booklets are not available on the Internet anymore.

'English'. Some studies involved pictorial stimuli. In this case, the language used when giving the instructions was taken as the clue for the annotations. Finally, some studies did not mention the language under investigation or include any examples, such that it was impossible to determine which language was investigated. Such studies were annotated as 'NA'.

The annotation process was performed by two different coders and their consistency was double-checked. The inter-coder's annotations, based on a sample of 322 abstracts, were the same for 91.18% of the cases, suggesting that there was no major problem during the annotation process. The cases with disagreement between the two coders mainly concerned studies involving mathematical modeling.

The dataset was then further annotated for (a) the code of the language in the Glottolog database (Hammarström et al., 2022), (b) its language family, (c) whether it is an LOL language based on the list proposed by Dahl (2015), and (d) its geographic information in terms of latitude and longitude (data retrieved from the Glottolog database). The information concerning language families were divided into two kinds: *Top-level family* and *Language family*. Top-level family corresponds to the highest node in the family tree in the Glottolog database for a given language. However, this did not allow us to distinguish between different languages within the Indo-European family. Therefore, language family corresponds to (a) the highest node in the family tree for non-Indo-European languages, and (b) the second highest node for Indo-European languages (e.g., Germanic, Romance, Celtic, etc.).

4.2.2 Analyses

The analyses were divided into four kinds: (a) individual-language analysis, (b) typological analysis, (c) sociolinguistic analysis (corresponding to LOL and non-LOL languages), and (d) the analysis of specific linguistic phenomena.[7]

The individual-language analysis consists in identifying the total number of languages found in the studies, the ten most represented languages, and the evolution of the number of languages from 2012 to 2023. Comparisons were made between the three types of conferences concerning the total number of studies and the most represented languages, while the evolution of the number of languages only concerned the major conferences because these conferences provide data across the full range of years under investigation.

The typological analysis replicates the pipeline in Kidd and Garcia (2022). Namely, the languages were categorized into three types: English, Other Indo-European, and Non-Indo-European. This pipeline was adopted after observing that a large majority of the studies were conducted on English, such that including English in the Indo-European family for the analysis

---

[7] The software R (R Core Team, 2018) was used to conduct the analyses, and the ggplot2 package (Wickham, 2009) to draw the plots.

may have biased and blurred the overall picture. The number of studies per language group as well as their evolution over the years were included in the analyses. However, unlike Kidd and Garcia (2022), the proportion of studies per language group and their evolution from 2012 and 2023 were also taken into account, as the overall number of studies was different for each year and conference. The evolution over time was based on the three major conferences. In addition, the number and proportion of studies per language group were compared between the three types of conferences. The potential difference in proportion of studies between the language groups may be due to the availability of grammars describing the systems of these languages.[8] The number of grammars as well as their year of publication were retrieved from the Glottolog database by filtering two types of documents: (a) 'grammar' and (b) 'sketch grammar'. Only documents published after 1900 were included.

The third analysis focuses on the representation of LOL and non-LOL languages in the abstracts, which may reflect an under-documented sociolinguistic bias in language processing research. The same variables as previously described were considered, i.e., the number and proportion of studies conducted on LOL and non-LOL languages, as well as (a) their evolution over time (data from the three major conferences), and (b) their comparison between the three types of conferences.

The fourth analysis concerns the representativity of two linguistic phenomena: (a) specific grammatical features: morphosyntactic alignment, richness of case morphology, and basic word order (Bornkessel-Schlesewsky et al., 2011), and (b) the conceptualization of temporality.[9] These phenomena were selected because they are subject to significant cross-linguistic variation (Bhat, 1999; Givón, 2001; Van Valin, 2005). The relevant abstracts were first automatically retrieved with keywords and then manually checked.[10] Concerning specific grammatical features, each entry was annotated depending on the relevant characteristics of the languages under investigation: (a) alignment type of the language (accusative, ergative, split ergative, symmetrical, other types of alignment), (b) the richness of morphological case marking, and (c) basic word order (Bornkessel-Schlesewsky et al., 2011). Basic word order was further annotated according to the

---

[8] This idea was advanced by several linguists when discussing the conceptualization of the present study.

[9] Brief explanatory notes on these phenomena are given in Appendix A. Unlike Kidd and Garcia (2022), we did not annotate the corpus in terms of phonology, morphosyntax, vocabulary and semantics, pragmatics and discourse, language and cognition, and non-verbal communication, because drawing a clear-cut line between these different domains is not easy, and this depends a lot on the linguistic theory we rely on (Butler, 2003, Vol. 1, Chapter 2, pp. 22–62; Francis, 2022, pp. 18–42). Indeed, some linguistic phenomena may be characterized as syntactic by some theories, whereas others may rather consider them as semantic or at the syntax-semantics interface.

[10] The keywords were: (a) specific grammatical features: thematic relation, thematic role, argument structure, Actor, Undergoer, Agent, Patient, alignment, accusative, ergative, symmetrical, Austronesian, word order, canonical order, SVO, SOV, VSO, VOS, OSV, OVS, verb-initial, verb-final, subject-initial, subject-final, object-initial, object-final, and scrambl(e/ing); (b) temporal concepts: tense, aspect, mood, modality, temporal concord, time reference, temporal reference.

canonical order between the verb and the subject, the verb and the object, and the subject and the object. For temporal concept processing, these were (a) the topic of the research question, and (b) the primary meaning of the grammatical temporal marker (tense, aspect, modality or mood) as described in the studies. These four analyses are summarized in **Figure 1**.
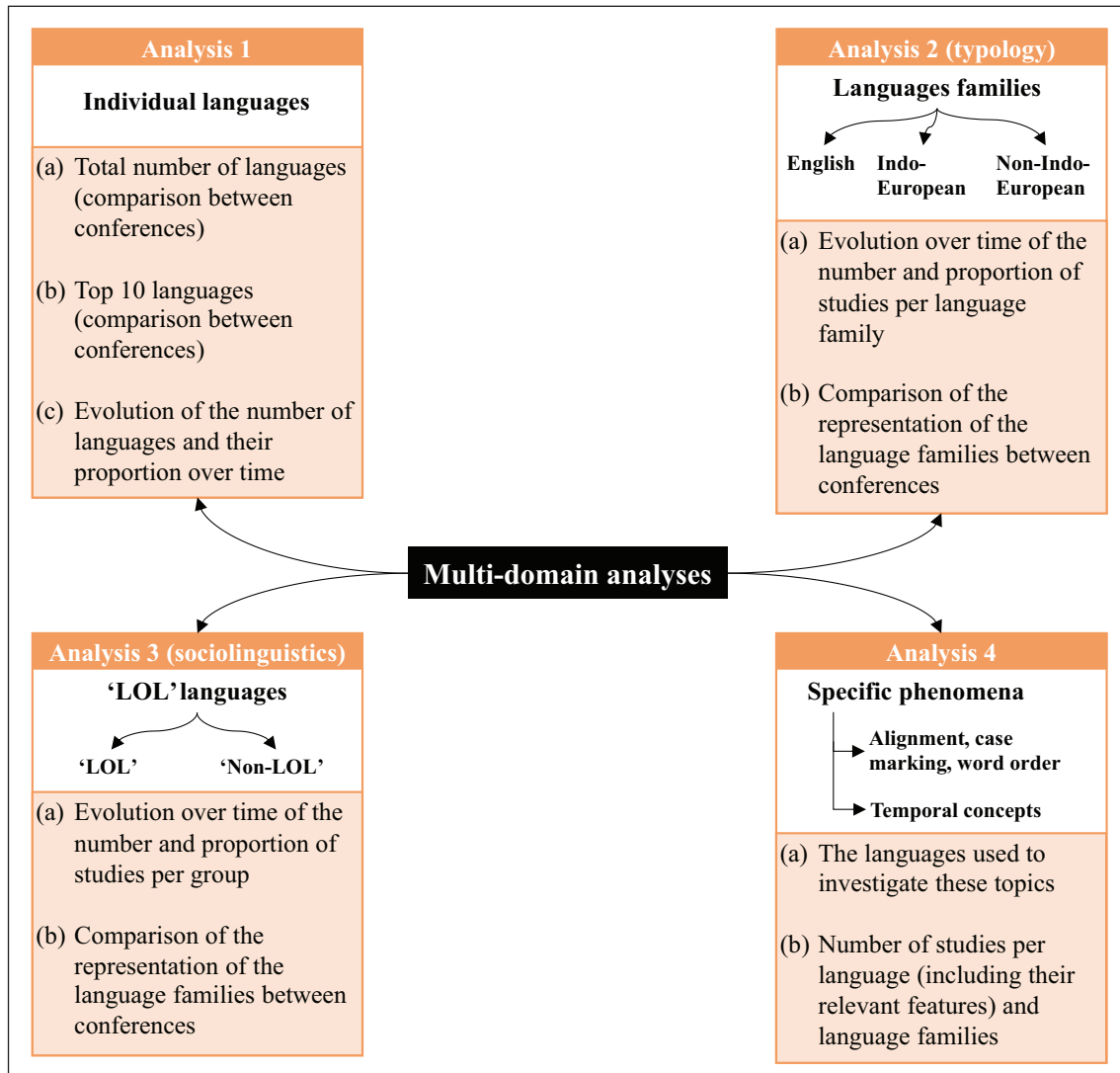

Analysis 1

Individual languages

(a) Total number of languages (comparison between conferences)

(b) Top 10 languages (comparison between conferences)

(c) Evolution of the number of languages and their proportion over time

Analysis 2 (typology)

Languages families

English   Indo-European   Non-Indo-European

(a) Evolution over time of the number and proportion of studies per language family

(b) Comparison of the representation of the language families between conferences

Multi-domain analyses

Analysis 3 (sociolinguistics)

'LOL' languages

'LOL'   'Non-LOL'

(a) Evolution over time of the number and proportion of studies per group

(b) Comparison of the representation of the language families between conferences

Analysis 4

Specific phenomena

→ **Alignment, case marking, word order**

→ **Temporal concepts**

(a) The languages used to investigate these topics

(b) Number of studies per language (including their relevant features) and language families

Figure 1: Summary of the analyses.

The evolution over time of the distribution of the languages and language families was statistically analyzed with Poisson regression models using the *glm* function from the lme4 package (D. Bates et al., 2015). We chose to closely replicate the statistical analysis pipeline in Kidd and Garcia (2022) for ease of comparisons with their findings. When comparing the distribution of the studies in terms of language families, the English group was taken as the reference level for the

(Other) Indo-European and Non-Indo-European groups. As for the comparison of this distribution in terms of LOL languages, the LOL Languages group was taken as the level of reference for the Non-LOL Languages group (dummy coding was used for the models). The value of the years was adjusted such that 2012 corresponded to 'Year 0', hence playing the role of the year of reference to track the changes over time. The dependent variables were (a) the number of languages under investigation, or (b) the proportion of studies by year, depending on the nature of the analysis, as previously described. The models were selected based on the AICc value of different models with the aictab function of the AICcmodavg package (Mazerolle, 2020). The model with the lowest AICc (a version of AIC corrected for smaller sample size) was selected for further analysis. This selection was further checked with base R ANOVAs, using likelihood ratio tests and selecting more complex models when the fit to the data was significantly improved.[11] The competing models consisted of (a) the model without any factor (i.e., intercept-only model), (b) the model with only the factor of language/language subgroup/LOL language, (c) the model with only the factor of year, (d) the model including both factors, and (e) the model including the interaction of both factors. The data observed in the corpus were then plotted in addition to the predicted data from the models.

4.3 Results

4.3.1 Analysis 1: Individual languages

A total of 8403 studies were included in the corpus after the annotation process for 9497 datapoints.[12] These studies covered in total 139 languages, which is higher than the number reported by Kidd and Garcia (2022). 120 languages were represented in the major conferences, 33 in the Asian ones, and 37 in the cross-linguistic ones. The ten most represented languages covered 86.07% of the studies overall. While this proportion was similar in the major conferences (86.97%) and the Asian conferences (88.61%), the situation was different for the cross-linguistic conferences, as the top 10 languages accounted for 56.58% of the studies.[13] See Appendix B for the number and proportion of studies for each language and each type of conference.

The abstracts are highly skewed towards English in particular (53.23% of the studies in the major conferences), as well as other Indo-European languages (28.77% of the studies in the major conferences). This is, again, similar to the results in Kidd and Garcia (2022). Is this skewness the same over the years? The number of languages investigated in conferences per year as well as the predicted trend from linear regression are shown in **Figure 2**.

---

[11] The two methods used to select the models gave consistent results.

[12] The difference between the number of studies and the number of datapoints is due to the fact that some studies investigated more than one language. The percentages refer to the total number of datapoints.

[13] It must be noted that this arbitrary cut is not suited for the cross-linguistic conferences, as it does not include groups of languages with the same number of studies. If we take languages with at least two studies, the top 17 languages cover 73.68% of the studies.
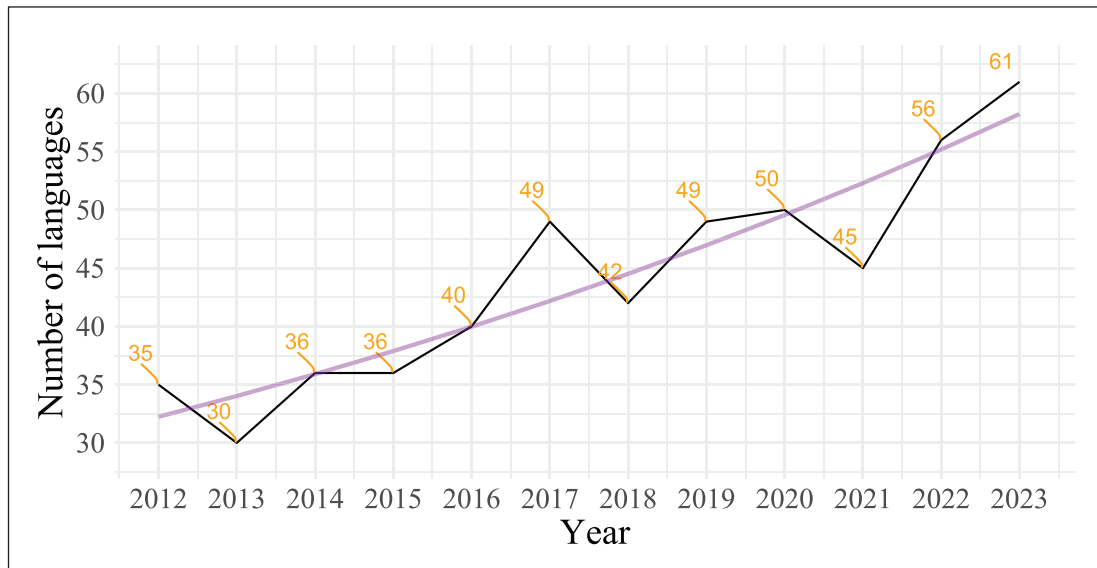
Figure 2: Number of languages investigated in conferences per year (black line: observed data, purple line: predicted data).

The statistical models (summarized in **Table 4**) indicate that the main effect of Year on the number of languages found in conference abstracts is highly significant ($\beta = 0.05$, $SE = 0.01$, $z = 4.2$, $p < .0001$). As **Figure 2** indicates, the number of languages grew linearly over the years, even if a small portion of them covers the majority of studies.

Finally, we note that this skewness towards one language differs from one conference to another. In the Asian conferences, 43.61% of the studies were conducted on Mandarin Chinese, when the second language is English (15.00%). As for the cross-linguistic conferences, even if English is also the most represented languages, its proportion is highly reduced (10.53% of the studies).

4.3.2 Analysis 2: Representation of language families

Not only is there skew towards just a few well-studied languages, but these languages are themselves not very diverse in terms of typology (see Appendix B, notably concerning the major conferences). In the major conferences, seven of the top 10 languages belong to the Indo-European family, one to Sino-Tibetan, one to Japonic and one to Koreanic. Our next question is whether the number and proportion of studies of different language families changed over the years. In other words, is the trend depicted in **Table B.1** (Appendix B), that English and Indo-European languages represent the majority of the studies, stable, or can we observe any changes over time?

The number of studies per year focusing on English, other Indo-European languages and languages from other families is given in **Figure 3**, panel A, and their percentage in panel B. The best-fitting model for the number of studies included the interaction between the two fixed

factors, i.e., Year and Language family (see **Table 4** for a summary of the results of the statistical models). Concerning the number of studies per language family per year, the highly significant interactions between (a) Year and Other Indo-European languages, and (b) Year and Other languages suggest that while the number of studies grew over the years, they tended to grow faster for these language groups than English, as indicated by panel A in **Figure 3**.

Table 4: Summary of the regression models on the number and proportion of language families over the years.

Analysis	Sub-analysis	Predictors	Estimate	*SE*	*z* value	*p* value
Individual languages	**Number of studies**	Intercept	3.47	0.09	38.86	<.0001
		Year	0.05	0.01	4.22	<.0001
Language families	**Number of studies**	Intercept	5.90	0.03	211.97	<.0001
		Year	0.02	0.00	4.11	<.0001
		Indo-European	–0.86	0.05	–17.46	<.0001
		Other	–1.47	0.06	–24.43	<.0001
		Year * Indo-European	0.04	0.01	5.78	<.0001
		Year * Other	0.06	0.01	7.55	<.0001
	Proportion	Intercept	4.12	0.07	57.97	<.0001
		Year	-0.02	0.01	–2.18	<.05
		Indo-European	–0.91	0.13	–7.12	<.0001
		Other	–1.50	0.16	–9.60	<.0001
		Year * Indo-European	0.05	0.02	2.43	<.05
		Year * Other	0.07	0.02	2.84	<.01

The results concerning the proportion of studies focusing on these three groups of languages depict a similar trend. Again, the best-fitting model included the interaction between Year and Language subgroup. The results of the regression models given in **Table 4** further show that the interaction between Year and Other Indo-European languages, as well as the interaction between Year and Other languages, were significant (all *p*'s below 0.05). Along with visual inspection of the results in panel B of **Figure 3**, these indicate that the proportion of studies focusing on English decreased over the years in favor of other languages, Indo-European and non-Indo-European. It may even be remarked that the proportion of English psycholinguistic and neurolinguistic studies was below 50% in 2023, the most recent year in the present survey.

Differences between the types of conference must be noted. The proportion of non-Indo-European studies was the highest in Asian conferences (69.17%) and cross-linguistic conferences (51.32%), while English studies were lower (15.00% and 10.53%, respectively).
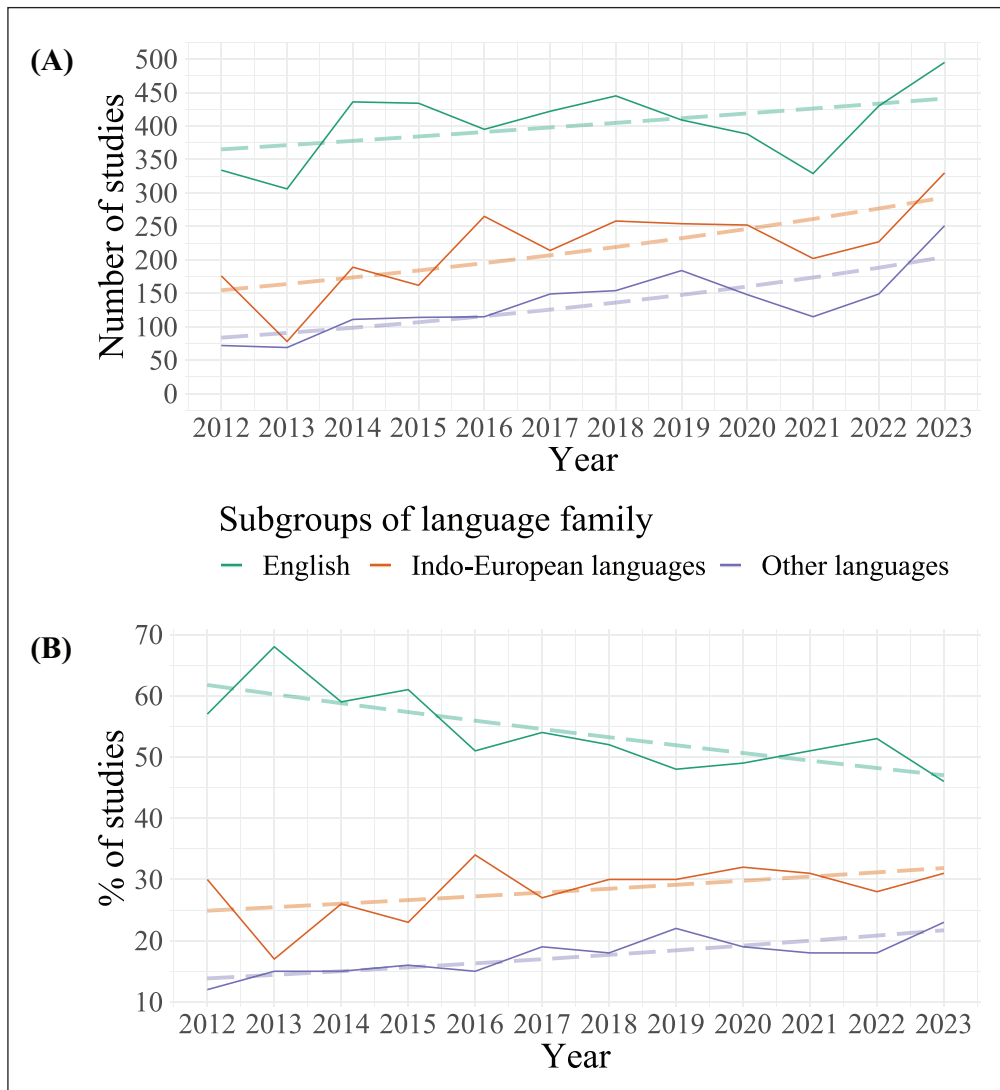
Figure 3: Investigated language groups in the abstracts of major conferences per year. Panel A: number of studies, Panel B: percentage of the total number of studies). Solid line: corpus observations, dashed line: regression line of predicted data.

The proportion of Indo-European (non-English) studies was 38.16% in the cross-linguistic conferences, and only 15.83% in the Asian conferences. A closer look at these distributions shows that more than 81.51% of the Indo-European studies concern Romance and Germanic languages in the major conferences, while 84.21% were Indo-Iranian languages in the Asian conferences, and these proportions were more balanced in the cross-linguistic conferences. As for non-Indo-European languages, the studies were highly skewed towards the Sino-Tibetan family (41.32% and 69.48% in the major and Asian conferences, respectively), especially Mandarin Chinese, while the proportions were again more balanced in the cross-linguistic conferences.

4.3.3 Representation of LOL/non-LOL languages

In addition to addressing linguistic diversity in conference abstracts in terms of language families, another question concerns the representativity of sociolinguistically dominant vs. non-dominant languages, here reflected by LOL vs. non-LOL languages, based on Dahl's (2015) criteria. 45 out of the 57 LOL languages identified by Dahl (2015) were found in the abstracts of major conferences, while there were 93 non-LOL languages.[14] The 45 LOL languages covered 8706 studies, or 96.08% of the abstracts, and the non-LOL languages, 355 studies (3.92% of the abstracts).

Can we observe changes over time concerning LOL and non-LOL languages? The statistical results of the change across years in the numbers and proportions of LOL and non-LOL language studies are given in **Table 5**. The best-fitting model for the number of studies included the interaction between Year and Number of LOL-language studies ($\beta = 0.03$, $SE = 0.02$, $z = 2.06$, $p < .05$). Visual inspection of the results in **Figure 4** (panel A) suggests that while studies of LOL and non-LOL languages were more and more numerous over time, the growth of non-LOL languages was more rapid than LOL languages, as they more than doubled over time.

Table 5: Summary of the regression models on the number and proportion of LOL-language studies over the years.

Analysis	Sub-analysis	Predictors	Estimate	*SE*	*z* value	*p* value
LOL languages	Number of studies	Intercept	6.36	0.02	296.26	<.0001
		Year	0.04	0.00	12.49	<.0001
		LOL languages	–3.40	0.11	–29.66	<.0001
		Year * LOL languages	0.03	0.02	2.06	<.05
	Proportion	Intercept	4.56	0.03	154.44	<.0001
		LOL languages	–3.33	0.16	–20.96	<.0001

The best-fitting model for the proportion of studies of LOL and non-LOL languages only included the main factor of LOL languages, which was highly significant. This and panel B of **Figure 4** suggest that the proportions of LOL languages were much higher than non-LOL languages, without significant changes over time (between 2.60% and 5.40% of the total number of studies).

Again, differences between types of conferences can be observed. In the major conferences, the proportion of LOL language studies was 96.08% (8706 out of 9061 studies), which is similar to the proportion in Asian conferences (88.33% of LOL language studies, corresponding to 318

---

[14] The LOL languages absent in the present survey are Afrikaans, Amharic, Belarusan, Eastern Panjabi, Gujarati, Halh Mongolian, Kazakh, Kyrgyz, Latvian, Malagasy, Sinhala, Swahili, Ukrainian, and Yoruba.

out of 360 studies). However, this proportion is reduced for cross-linguistic conferences, even if LOL languages studies are still dominant: 73.68% (56 out of 76 studies).
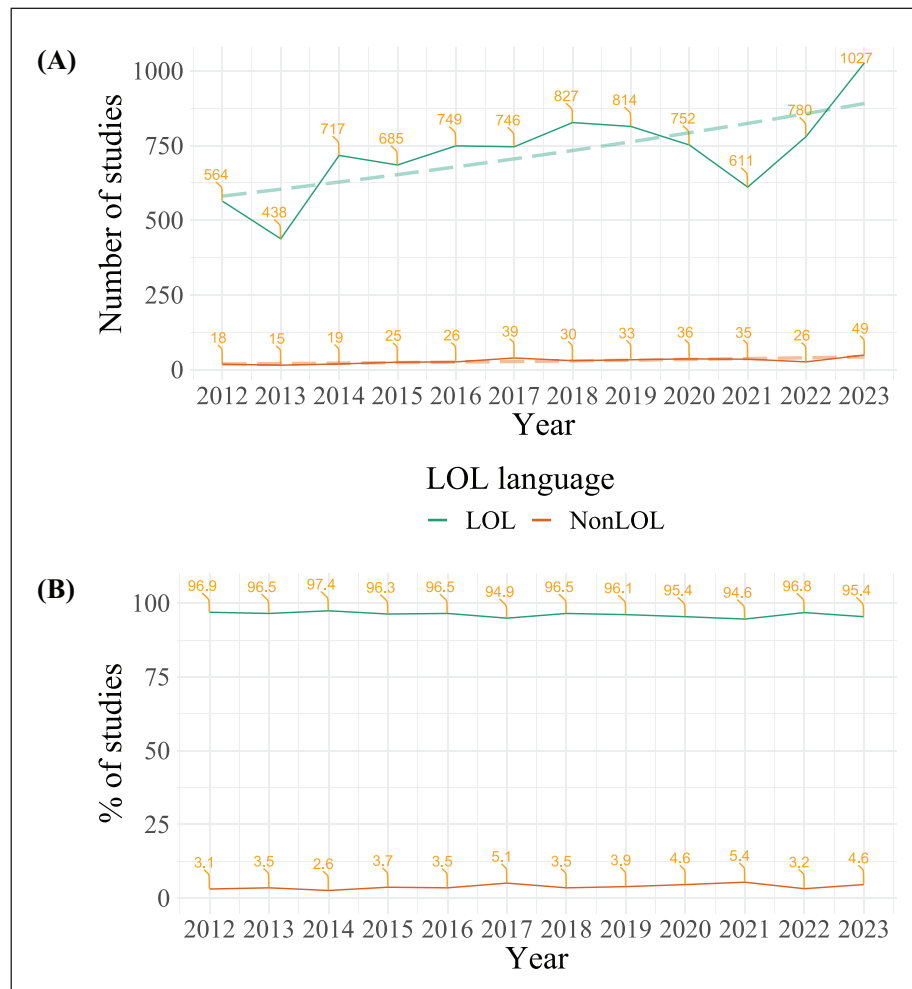

Figure 4: Distribution of the LOL and non-LOL languages in the abstracts of major conferences per year. Panel A: number of studies, Panel B: percentage of the total number of studies). Solid line: corpus observations, dashed line: regression line of predicted data.

4.3.4 Analysis 4: Specific phenomena

4.3.4.1 Specific grammatical features: Alignment type, richness of case marking and basic word order

In total, 375 studies related to language-dependent grammatical features (Givón, 2001; Van Valin, 2005) were retrieved. The proportions of studies focusing on (a) the alignment type of the language, (b) its richness of morphological case marking, and (c) its basic word order, according to subgroups of language families, are given in **Tables 6**, **7** and **8**.

Table 6: Number and percentage of studies related to the morphosyntactic alignment of the language.

	Accusative	Ergative	Split ergative	Symmetrical	Other types	*Total*
English	120 (32%)	/	/	/	/	*120 (32%)*
Indo-European	108 (28.8%)	0	14 (3.7%)	/	0	*122 (32.5%)*
Other	89 (23.7%)	20 (5.3%)	5 (1.3%)	12 (3.2%)	7 (1.9%)	*133 (35.5%)*
Total	*317 (84.5%)*	*20 (5.3%)*	*19 (5.1%)*	*12 (3.2%)*	*7 (1.9%)*	*375 (100%)*

Table 7: Number and percentage of studies related to the richness of morphological case marking of the language.

	Poor	Rich	*Total*
English	120 (32%)	/	*120 (32%)*
Indo-European	45 (12%)	77 (20.5%)	*122 (32.5%)*
Other	66 (17.6%)	67 (17.9%)	*133 (35.5%)*
Total	*231 (61.6%)*	*144 (38.4%)*	*375 (100%)*

Table 8: Number and percentage of studies related to the basic word order of the language.

	SVO	SOV	SVO/SOV	VSO	VOS	OVS	*Total*
English	120 (32%)	/	/	/	/	/	*120 (32%)*
Indo-European	41 (10.9%)	18 (4.8%)	62 (16.5%)	1 (0.3%)	0	0	*122 (32.5%)*
Other	43 (11.5%)	64 (17.1%)	0	18 (4.8%)	7 (1.9%)	1 (0.3%)	*133 (35.5%)*
Total	*204 (54.4%)*	*82 (21.9%)*	*62 (16.5%)*	*19 (5.1%)*	*7 (1.9%)*	*1 (0.3%)*	*375 (100%)*

The results in terms of morphosyntactic alignment show that the number of studies is skewed towards languages exhibiting an accusative alignment type, in particular, because of English (32%) and other Indo-European languages (28.8%) representing more than half of the studies. While the situation is more balanced regarding the richness of morphological case marking, we

can still observe that the relatively higher proportion of languages with a poor morphological case marking system is mostly due to the higher proportion of English studies (32%). Finally, there is a clear bias towards languages exhibiting an SVO and/or SOV basic word order. A similar bias can also be found when looking at other types of word orders, as illustrated in **Table 9**.

Table 9: Distribution of the basic order between V and S, V and O, and S and O according to language subgroupings.

	Order between V and S		Order between V and O			Order between S and O	
	V-S	*S-V*	*V-O*	*O-V*	*Both*	*S-O*	*O-S*
English	/	120 (32%)	120 (32%)	/	/	120 (32%)	/
Indo-European	1 (0.3%)	121 (32.3%)	42 (11.2%)	18 (4.8%)	62 (16.5%)	122 (32.5%)	0
Other	26 (6.9%)	107 (28.5%)	68 (18.1%)	65 (17.3%)	0	125 (33.3%)	8 (2.1%)
Total	*27 (7.2%)*	*348 (92.8%)*	*230 (61.3%)*	*83 (22.1%)*	*62 (16.5%)*	*367 (97.9%)*	*8 (2.1%)*

When crossing these three grammatical features, there are theoretically 60 potential configurations, and 19 of them were found in the investigation of the abstracts. They are summarized in **Table 10**. This crossing reflects the observations made above: the data were skewed towards languages exhibiting an accusative alignment system, an SVO and/or SOV word order, and, to a lesser extent, a poor morphological case system.

4.3.4.2 Temporal concepts

A total of 111 studies were retrieved which covered seven topics: (a) time reference, (b) lifetime effect, (c) event perception, (d) aspect processing, (e) aspectual coercion, (f) morphological form, and (g) time marking in specific syntactic constructions. We focus on the topics in (a–d) in the following, as they are more likely to exhibit cross-linguistic variation (total: 71 studies). Their distribution according to research topic and primary meaning of the temporal marker is given in **Table 11**.

We can see that the studies on time reference are biased towards languages exhibiting tense marking (12 out of 22 studies), and to a lesser extent, aspect marking (7 studies). Similar remarks can be made for lifetime effect studies, which have never been investigated from the angle of modality and mood marking.[15]

---

[15] The event perception and aspect processing studies only involve aspect marking because of their nature.

Table 10: Number and percentage of studies related to (a) the alignment type of the language under investigation, (b) its richness of morphological case marking, and (c) its basic word order.

	Alignment type									
	Accusative		*Ergative*		*Split ergative*		*Symmetrical*		*Other*	
	Case morphology		**Case morphology**		**Case morphology**		**Case morphology**		**Case morphology**	
	Poor	*Rich*	*Poor*	*Rich*	*Poor*	*Rich*	*Poor*	*Rich*	*Poor*	*Rich*
SVO	184 (49.1%)	15 (4%)	0	0	0	0	0	0	5 (1.3%)	0
SOV	2 (0.5%)	48 (12.8%)	0	13 (3.5%)	2 (0.5%)	16 (4.3%)	0	0	1 (0.3%)	0
SVO/ SOV	13 (3.5%)	49 (13.1%)	0	0	0	0	0	0	0	0
VSO	4 (1.1%)	1 (0.3%)	1 (0.3%)	0	1 (0.3%)	0	11 (3%)	0	0	0
VOS	0	0	4 (1.1%)	2 (0.5%)	0	0	1 (0.3%)	0	0	0
OVS	1 (0.3%)	0	0	0	0	0	0	0	0	0

Table 11. Distribution of research topics and meanings of the temporal markers in the studies on temporal concepts processing.

	Time reference	Lifetime effect	Event perception	Aspect processing
Tense	12 (54.55%)	1 (25%)	0	0
Aspect	7 (31.82%)	3 (75%)	11 (100%)	34 (100%)
Modality	2 (9.09%)	0	0	0
Mood	1 (4.55%)	0	0	0

5. Discussion

5.1 Linguistic diversity and language processing: Where are we now?

5.1.1 Language processing research and typological bias: The big picture

The present paper surveyed the degree of linguistic diversity in international psycholinguistic and neurolinguistic conferences over the past decade. As Anand et al. (2011) point out, these two fields are biased towards a small group of languages, mainly English and other Indo-European languages. The survey of language acquisition studies conducted by Kidd and Garcia (2022) draws similar conclusions, i.e., language acquisition studies focused primarily on Indo-European languages, but the number of articles on Indo-European languages and other languages grew

more rapidly than the number of articles on English over time. Following these two studies, the present article surveyed linguistic diversity in seven language processing conferences from 2012 to 2023 by looking at different factors: (a) the number of languages found in the abstracts, (b) the number of studies and their proportion according to three groups of languages – English, other Indo-European and non-Indo-European languages (i.e., the potential typological bias), (c) the number of studies and their proportion in terms of LOL and non-LOL languages (i.e., the potential sociolinguistic bias), and (d) the linguistic diversity of specific linguistic phenomena. The results are summarized in **Table 12**.

Table 12: Summary of the results of the present survey.

Analysis	Sub-analysis	Replication of analysis pipeline	Results
Number of languages	Overall number	Anand et al. (2011), Kidd & Garcia (2022)	139 languages (120 in the major conferences)
	Over time	Kidd & Garcia (2022)	Increased diversity
Studies according to language families	Number of studies	Kidd & Garcia (2022)	Increase for all language groups, more rapidly for Indo-European (except English) and non-Indo-European languages
	Proportion of the number of studies	/	Decrease for English, increase for other Indo-European and non-Indo-European
Studies according to LOL languages	Number of studies	/	Increase for both groups, more rapidly for non-LOL languages
	Proportion of the number of studies	/	No change over time, proportion of LOL-language studies above 94%
Specific phenomena	Morphosyntactic alignment, richness of case morphology, basic word order	/	Bias towards languages with accusative alignment, SVO and/ or SOV word order and, to some extent, poorer case morphology
	Temporal concepts	/	Bias towards language with grammatical tense marking

Overall, the results of the present survey indicate that even if the total number of represented languages is quite small compared with the number of languages in the world, language processing research presented at conferences is becoming more and more diversified. It must be remarked

that the number of languages under investigation linearly increased over time: 35 in 2012, and 61 in 2023. Nevertheless, this finding hides another observation, i.e., the large majority of these studies (86.07%) only covers ten languages. Even so, as the overall number of psycholinguistic and neurolinguistic studies increased in the last decade, the number of abstracts focusing on a language different from English (Indo-European and non-Indo-European) also increased more quickly over time. This is even more clear if we look at the proportion of studies focusing on English, which shows a tendency to decrease in favor of other languages. This is another indication that psycholinguistic and neurolinguistic research is becoming more and more diversified.

The data concerning LOL and non-LOL languages depict another picture, as their representation in terms of proportion did not significantly change over time, and the proportion of abstracts focusing on LOL languages is much higher than non-LOL language, exceeding 94%. Nevertheless, we note that the number of studies on both LOL and non-LOL languages increased over the last decade, and the pace of increase was slightly quicker for non-LOL than LOL languages. This may only be an effect of the increase in the overall number of studies in general.

This bias may be due to the nature of the major conferences, mainly held in Western countries and therefore perhaps more accessible to researchers working in the areas where Indo-European languages are predominantly spoken. The results from the Asian conferences show that the languages which are more commonly investigated are indeed languages spoken in this area, mainly Mandarin Chinese, Hindi, Japanese and Korean, reflecting the effect of the geographical location of the conference. Nevertheless, linguistic diversity in the Asian conferences is not greater: the top five languages account for 80.27% of the presented studies. Conferences focusing on cross-linguistic research in language processing seem to exhibit greater linguistic diversity. However, the overall number of studies presented at these conferences is still rather low compared with the major and Asian venues, and more observations should be collected to determine whether this pattern is stable over time.

In sum, language processing research is becoming more and more diversified. But this trend is not exactly the same as what is observed in language acquisition research. According to Kidd and Garcia (2022), the number of non-English Indo-European studies also grew faster than English studies (but English still predominates), yet the rate of growth of language acquisition studies appears to be faster than for general language processing. Their study indicates that there have been more Indo-European language acquisition studies than English ones by the end of the last decade, while English studies still outnumber other languages in our survey. This may be a consequence of the fact that, according to Kidd and Garcia (2022, p. 705), "[language acquisition research] has a justifiably proud tradition of cross-linguistic work". The overrepresentation of languages which are typologically related is not trivial when it comes to language processing, since they are likely to share similar features (Haspelmath, 2001), even if it is true that diversity can be found within the Indo-European language family (for example, German exhibits systematic

case marking and well-developed verbal inflection in comparison with English). We discuss this point in 5.2.

5.1.2 LOL languages: Additional sociolinguistic bias?

The typological bias discussed in the previous subsection is not only inherent to language processing research, but also to theoretical linguistics, even though this has changed over time. Dahl (2015) proposed that in linguistics, this sociolinguistic bias has now overtaken the typological one: languages which are more frequently studied are more likely to be official languages with lots of speakers and a writing system used by the majority of them. The data in our survey indicate that the same sociolinguistic bias is found in language processing research, with only minor changes over time, as the report of the number and proportion of LOL and non-LOL languages suggests. This sociolinguistic bias, among other, can be seen as a consequence of the challenges that researchers face when attempting to conduct experiments on less studied languages. We discuss this point in more detail in 5.3.2.

5.2 Linguistic representativity of specific phenomena

The results from the analysis of the particular phenomena presented in 4.3.4 show that the processing of certain grammatical configurations has been investigated in many studies, while others have not been explored so far. In this subsection, we further discuss the impact of typological bias on these observations, and we will show how the inclusion of non-Indo-European languages is crucial for some situations.

5.2.1 Morphosyntactic alignment, richness of case morphological system and basic word order

The results in **Tables 6** to **10** indicate that there is a clear bias towards languages exhibiting (a) an accusative-type alignment system, and (b) an SVO and/or SOV basic word order. As we show in the following, these biases are also correlated with the typological bias found in the abstracts. We focus our discussion on morphosyntactic alignment systems and basic word orders.

In terms of typology, English accounts for one third of the entire data set, thus skewing the results towards accusative alignment and SVO word order. The analysis of the language families represented in conference abstracts indicated that in addition to English, there is also an overrepresentation of the Indo-European language family. This family is diverse when it comes to morphosyntactic alignment and basic word order, including, for example, languages with a split ergative system (e.g., Hindi) or VSO basic word order (e.g., Irish). However, the data in **Tables 6** and **8** show that despite this family-internal diversity, most of the Indo-European languages represented in the present survey exhibit an accusative alignment system (108 studies

out of the 122 non-English Indo-European studies, accounting for 88.52%), as well as an SVO and/or SOV basic word order (121 studies out of 122, or 99.18%). In total, 58 studies focusing on a language with an alignment system other than accusative were retrieved. Among them, 14 studies involved an Indo-European language, and 44, a non-Indo-European language. Crucially, the non-accusative Indo-European language studies only covered the split ergative alignment system, while the other languages focused on a larger variety of systems, including ergative (e.g., Basque and Tzeltal), split ergative (e.g., Tongan), so-called symmetrical (e.g., Tagalog and Truku Seediq) as well as other types of alignment systems (e.g., Zapotec).

This typological bias towards accusative languages actually correlates with the sociolinguistic bias discussed in the previous subsection. Indeed, only 12 of the 317 studies involving an accusative language included a non-LOL language. In contrast, the representativity of non-LOL languages is higher in studies with a language exhibiting a different alignment system: 32 out of 58.

The same typological bias occurs concerning the basic word order of the languages found in the abstracts, as most of the studies involved a language with basic SVO and/or SOV word order (348 studies out of 375, or 92.80%). Among them, English accounts for 120 studies, other Indo-European languages, for 121 studies, and non-Indo-European languages, for 107 studies. Again, grammatical variety can be found within the Indo-European language family. But in our survey, only one language has been the focus of a single study on an Indo-European language with a different basic word order: Irish.

Another interesting observation concerns the contribution of the non-Indo-European studies of canonical word order, which in total account for 35.50% of these studies. First, we note that the majority also display either SVO or SOV word order (107 studies out of 133, or 80.45%), reflecting the typological observation that these are the most common word orders (Dryer, 2013). However, more variety can be found within this group. For example, where only one instance of a VSO-language study was found for the Indo-European languages (Irish), there are 18 instances within the non-Indo-European group, covering five languages. There are also instances of word order which were not represented in the Indo-European group in our survey: VOS (Kaqchikel, Truku Seediq and Tzeltal) and OVS (Äiwoo).

We can also observe a sociolinguistic bias concerning the distribution of the studies according to the basic word order of the languages under investigation. Indeed, with regard to the 348 studies involving a language exhibiting one of the most represented word orders (i.e., SVO and/or SOV) in our survey, 316 (or 90.80%) include a LOL language, and only 32 (or 9.20%), a non-LOL language. By contrast, when it comes to the other 27 studies representing different word orders (i.e., VOS, VSO and OVS), 15 (or 55.56%) involve a LOL language, and 12 (or 44.44%), a non-LOL language.

All in all, the investigation of specific grammatical phenomena subject to cross-linguistic differences leads to similar conclusions as the big picture analysis, in that the typological bias towards

English and other Indo-European languages is quite pervasive. In addition to this typological bias, the sociolinguistic bias reflected by the distribution of LOL and non-LOL languages can also be found in the case of phenomena which are mostly represented by English and other Indo-European languages, indicating that there are indeed two layers of bias in language processing research.

5.2.2 Temporal concepts

The analysis of the processing of temporal concepts revealed that there was a bias towards tense marking (and to a lesser extent, grammatical aspect marking) concerning time reference processing. The breakdown of these studies in terms of language group and the category of the grammatical marking under investigation is given in **Table 13**.

Table 13: Number of studies according to the grammatical temporal concepts and the subgroup of languages under investigation.

English	Indo-European	Non-Indo-European			
Tense	*Tense*	*Tense*	*Aspect*	*Modality*	*Mood*
1	8	3	7	2	1

The data in **Table 13** suggest that the Indo-European bias can also be seen in studies of time processing, since together, Indo-European languages account for 40.91% of the data. In terms of grammatical marking of time reference, Indo-European languages more often rely on tense marking and can even be considered 'tense-prominent' (Bhat, 1999; Comrie, 1985). Therefore, the general typological bias can again be observed concerning the particular phenomena of time reference processing. The inclusion of non-Indo-European languages appears to be crucial: as illustrated by the present survey, the investigation of time reference processing with aspect, modality and mood markers relies exclusively on these languages. This is especially the case for modality and mood marking studies, which are clearly underrepresented, while such marking is not rare in the languages of the world, such as Burmese (Sino-Tibetan; Comrie, 1985, p. 51), Q'eqchi' (Mayan; Bhat, 1999, pp.132–133) as well as some Formosan languages (Austronesian; Zeitoun et al., 1996).

5.3 Challenges for the inclusion of non-Indo-European languages in language processing research

5.3.1 Some language processing studies do not necessarily require linguistic diversity

One of the reasons for the higher proportion of English (and, to some extent, Indo-European) studies may just be that not all research questions and objectives necessarily need linguistic

diversity. Indeed, as experimental tools and experimental paradigms are still developing, these need to be tested to ensure their validity. This is the case, for example, for online eye-tracking measurements implemented on the Ibex Farm/PCIbex platform (Drummond, 2013; Schwarz & Zehr, 2021). While using the eye-tracking technique remotely may lead to new possibilities concerning data collection, it is crucial to address the issue of its validity, notably by determining to what extent results obtained in the laboratory setting and remotely are similar (Kandel et al., 2022; Langlois et al., 2023). Even if preliminary results are promising, there is still no consensus about how data should be preprocessed and corrected, or about the degree of precision of linguistic effects which may be observed in both in-lab and remote eye-tracking. Consequently, it makes more sense to conduct such replication tests on languages with (a) prior well-established in-lab results and (b) many speakers available for dozens of experiments.

Another reason for choosing more commonly studied languages to conduct linguistic experiments is that the processing mechanisms related to some linguistic phenomena may not be completely clear at the moment, even in languages often encountered in psycholinguistics and neurolinguistics. Such explorations need a 'starting point', and as fewer prior studies are available to rely on, it can be considered safer to begin with languages with more speakers. Indeed, it is possible that some factors may not have been considered in such preliminary studies, and therefore, very often, follow-up investigations are needed in order to establish more confidently the processing patterns and factors at stake.

5.3.2 Language 'availability': Academic and sociolinguistic considerations

Other obstacles to linguistic diversity in language processing research involve geographical, sociolinguistic, and 'academic availability' considerations.

Including linguistic diversity (non-Indo-European languages, in particular) in language processing research very often entails investigating languages which are not spoken in geographic areas where the research institutions are. This is also the case concerning LOL languages, as the ones not represented in the survey are located outside of North America and Western Europe. Therefore, there is a geographic obstacle regarding the inclusion of non-Indo-European languages: conducting experimental research on these languages may require having a laboratory setting in another place, often abroad, and if no laboratory is available, moving the laboratory setting, which entails much more time, energy and uncertainty (Wagers & Chung, 2023). In addition, the speakers of many non-Indo-European languages, especially endangered ones, live far from urban areas and laboratories, such that the laboratory setting must be moved to nearby the participants' locations. This virtually rules out certain techniques, such as fMRI (functional magnetic resonance imaging) and MEG (magnetoencephalography), and entails much more preparation for EEG (electroencephalography) studies (Speed et al., 2018, pp. 201–202). Last

but not least, institutional (from home and/or host countries) and financial support are necessary for researchers involved in this enterprise in order to successfully drive such projects.

The sociolinguistic status of many non-Indo-European languages (with some being endangered) may also induce additional challenges, which have already been mentioned by Whalen and McDonough (2015), Speed et al. (2018), Wagers and Chung (2023) and Collart (2024).

The first challenge concerns the number of potential participants as well as the problem of the fluency of the speakers. Are there enough fluent native speakers? Are they literate in their language, meaning that if they are not, the experiments could only be in the auditory modality (entailing that the experimental material must be recorded by fluent native speakers and be prepared, which takes more time and energy than for the visual mode)? Are they also fluent in another language, increasing possible interference between the two languages? Are there young speakers, as most experiments involve participants whose age range is from 20 to 30 years old (Heinrich et al., 2010)? The issue of participants is also relevant for the recruitment process, since they are likely to belong to communities where the social norms and/or cultural values are not the same as in urban/institutional areas (Wagers & Chung, 2023, pp. 505–508). Recruitment strategies and communication patterns may differ greatly, and very often, psycholinguists and neurolinguists need to collaborate with native speakers and spend more time in the community to understand other social norms and cultural values and be accepted by the population.

The second challenge is related to ethical issues, especially for endangered languages, as the standards of institutional review boards, reflecting to some extent the cultural values of institutional urban areas, may be in contradiction with the social norms and cultural values of some language communities (Whalen & McDonough, 2015, pp. 14–15). This creates situations in which the administrative procedures to be completed before designing the experiments are unclear, and this creates further challenges to consider when conducting experiments on endangered non-Indo-European languages (Collart, 2024).

The third obstacle is linked to 'academic availability'. Indeed, conducting experimental linguistic research requires enough description and linguistic analyses of the language in question to be able to formulate an experimental research question, propose a solid design and prepare the linguistic material. This can take several forms. As previously mentioned, one consists in collaborating with native speakers who ideally have received linguistic training such that they can provide invaluable suggestions and help through each stage of the experimental study. But this presupposes that active linguistic research is done on the particular language, and it is not guaranteed that psycholinguists and neurolinguists willing to work on other languages can encounter such research groups at any time. This dilemma can be even more exacerbated in the absence of institutional and financial support.

Having enough linguistic descriptions also assumes that grammars and sketch grammars on many languages are available to psycholinguists and neurolinguists, i.e., that there is enough documentation. The data in **Figure 5** represent the number of grammars and sketch grammars of the languages found in the present corpus published yearly since 1900.
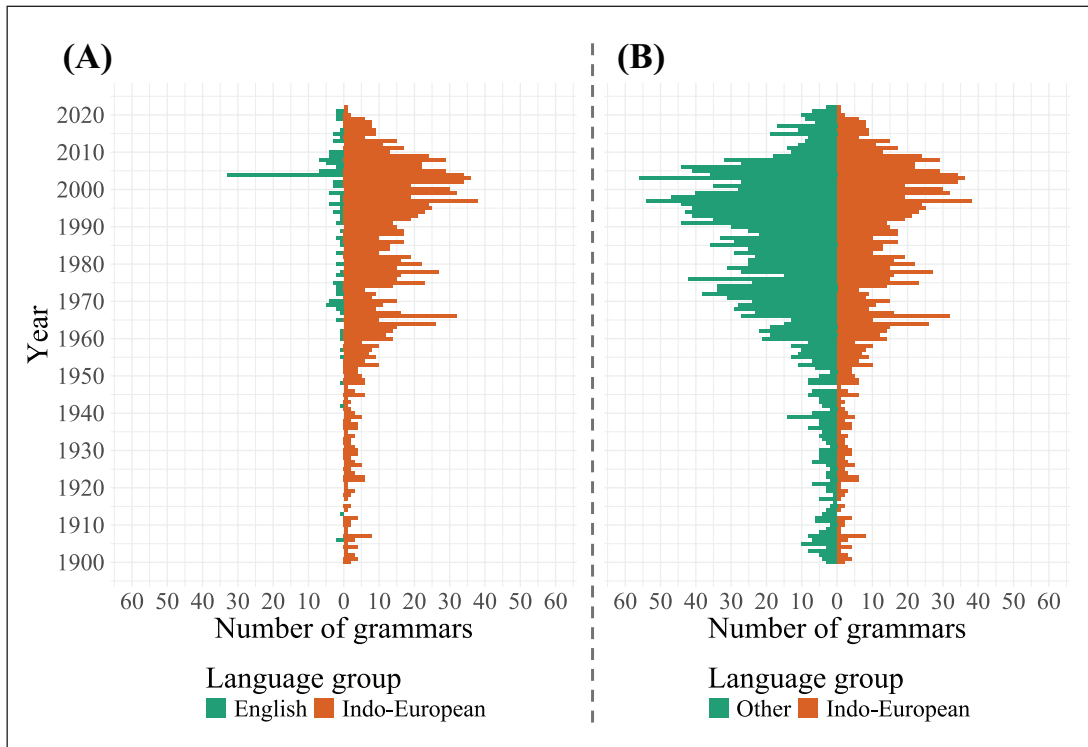

Figure 5: Number of grammars and sketch grammars of English, Indo-European and non-Indo-European languages published yearly since 1900. Panel A: comparison between English and other Indo-European languages; Panel B: comparison between Indo-European languages (English not included) and non-Indo-European languages.

The data in **Figure 5** show that there are more formal linguistic descriptions of non-Indo-European languages than Indo-European languages and English. These observations clearly contrast with the number and proportion of language processing studies on these three groups, as the reverse tendency was found in our survey. Therefore, while it is true that the availability of formal linguistic descriptions of a language is a necessary condition for conducting any experiment on it, we can see that this criterion does not explain the typological bias in the present survey. This may further suggest that the collaborative pattern between psycho/neurolinguists, linguists and native speakers described earlier in this section is more crucial when attempting the investigation of less commonly encountered languages.

6. Conclusion

The aim of this article was to survey linguistic diversity in psycholinguistic and neurolinguistic research by examining the languages under investigation in seven international conferences from 2012 to 2023. The results showed that these studies were highly skewed towards English and other Indo-European languages and that the number and proportion of Indo-European (other than English) and non-Indo-European languages increased over time, indicating that language processing research is getting more and more diversified linguistically. In addition, the number of non-LOL languages (based on the criteria proposed by Dahl, 2015), grew at a slightly faster rate than LOL languages, however their proportion did not change over time, reflecting the existence of a sociolinguistic bias in addition to a typological bias. This typological bias is found not only in the general picture, but also in several specific grammatical phenomena, such as the morphosyntactic alignment, richness of case morphology and basic word order of the languages under investigation, as well as the grammatical expression of temporal concepts. These results reflect the numerous challenges that we encounter when conducting experiments on typologically diverse and non-LOL languages, and the lack of their inclusion is likely to bias the generalization of processing patterns for at least some specific linguistic phenomena.

Appendix A. Notes on the definitions of specific phenomena in the fourth analysis

The term *morphosyntactic alignment* refers to the grammatical system used in a given language to distinguish between the arguments of the predicate (Dixon, 1994; Meyer, 2023; Van Valin, 2005; among many others).[16] Conventionally, three types of arguments are considered: (a) S, or the subject of an intransitive verb, (b) A, or the subject of a transitive verb (also referred to as the 'privileged syntactic argument' according to different theories), and (c) O, or the object of a transitive verb. Languages differ regarding the encoding of these arguments. *Accusative* languages encode S and A in the same way, while there is a distinguished treatment for O. In *ergative* languages, S and O are handled the same way, and a different treatment is reserved for A. *Split ergative* languages exhibit these two types of alignment depending on the specific grammatical construction (e.g., Hindi, where transitive verbs inflected for perfective aspect exhibit an ergative alignment, but an accusative one when inflected for imperfective aspect). *Symmetrical* (also sometimes referred to as 'Austronesian' or 'Philippine-type' in the literature) alignment refers to languages in which the relationship between the arguments and the predicate depends on its voice marking. For instance, verbs marked with 'Actor voice' will take the argument marked with nominative case (or other types of marking making it the 'privileged syntactic argument') as A, and the other ones as O, and vice-versa with 'Patient (or Undergoer) voice' marking.

The term *time reference* refers to the time frame in which the event is understood to happen (i.e., in the past, present or future). *Lifetime effect* refers to the grammatical means of referring to an individual's lifetime, and notably the incongruous use of present perfect marking with subjects referring to a dead individual (Klein, 1994; Musan, 1997). The term *event perception* refers to the degree to which our perception of how an event unfolds in time (i.e., about to start, ongoing, terminated, etc.) is affected by the use of time-related marking. The term *aspect processing* refers to the investigation of the processing of the morphosyntactic encoding of aspect.

Appendix B. Most represented languages in the major conferences, Asian conferences and cross-linguistic conferences

Table B.1. The ten most represented languages in the major conferences.

Rank	Language	Family	Number of studies	Percentage
1	English	Indo-European	4823	53.23
2	German	Indo-European	738	8.14
3	Mandarin	Sino-Tibetan	628	6.93

(Contd.)

---

[16] The explanations in Appendix A are intentionally simplified and the terms are used for descriptive purposes.

Rank	Language	Family	Number of studies	Percentage
4	Spanish	Indo-European	445	4.91
5	French	Indo-European	288	3.18
6	Dutch	Indo-European	286	3.16
7	Russian	Indo-European	213	2.35
8	Japanese	Japonic	156	1.72
9	Italian	Indo-European	154	1.70
10	Korean	Koreanic	149	1.64

Table B.2. The fourteen most represented languages in the Asian conferences.

Rank	Language	Family	Number of studies	Percentage
1	Mandarin	Sino-Tibetan	157	43.61
2	English	Indo-European	54	15
3	Hindi	Indo-European	35	9.72
4	Japanese	Japonic	30	8.33
5	Korean	Koreanic	14	3.89
6	Cantonese	Sino-Tibetan	13	3.61
7	Malayalam	Dravidian	6	1.67
8	Bangla	Austroasiatic	4	1.11
9	HKSL	Sign language	3	0.83
9	Assamese	Indo-European	3	0.83
9	Marathi	Indo-European	3	0.83
9	Telugu	Dravidian	3	0.83
9	TSL	Sign language	3	0.83
9	Vietnamese	Austroasiatic	3	0.83

Table B.3: The seventeen most represented languages in the cross-linguistic conferences.

Rank	Language	Family	Number of studies	Percentage
1	English	Indo-European	8	10.13
2	Mandarin	Sino-Tibetan	6	7.59
3	Spanish	Indo-European	5	6.33
4	Basque	Basque	4	5.06
4	Dutch	Indo-European	4	5.06
4	Hindi	Indo-European	4	5.06

Rank	Language	Family	Number of studies	Percentage
4	Turkish	Turkic	4	5.06
5	German	Indo-European	3	3.80
5	Hebrew	Afro-Asiatic	3	3.80
6	Chintang	Sino-Tibetan	2	2.53
6	Czech	Indo-European	2	2.53
6	Georgian	Kartvelian	2	2.53
6	Indonesian	Austronesian	2	2.53
6	Italian	Indo-European	2	2.53
6	Russian	Indo-European	2	2.53
6	Slavic	Indo-European	2	2.53
6	Tagalog	Austronesian	2	2.53

Data accessibility statement

The datasets (anonymized) and analysis codes are available at the following link: https://osf.io/ufk2r/.

Acknowledgements

Competing interests

The author has no competing interests to declare.

References

Anand, P., Chung, S., & Wagers, M. (2011). *Widening the net: Challenges for gathering linguistic data in the digital age*. White paper published in NSF project SBE 2020: Future research in the Social, Behavioral and Economics Sciences.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition, 11*(3), 245–299. DOI: https://doi.org/10.1016/0010-0277(82)90017-8

Bhat, D. N. S. (1999). *The prominence of tense, aspect and mood*. John Benjamins. DOI: https://doi.org/10.1075/slcs.49

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences, 26*(12), 1153–1170. DOI: https://doi.org/10.1016/j.tics.2022.09.015

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review, 113*(4), 787–821. DOI: https://doi.org/10.1037/0033-295X.113.4.787

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., Roehm, D., & Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain & Language, 117*(3), 133–152. DOI: https://doi.org/10.1016/j.bandl.2010.09.010

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). 11 Neurotopology: Modeling crosslinguistic similarities and differences in the neurocognition of language comprehension. In M. Sanz, I. Laka, & M. K. Tanenhaus (Eds.). *Language down the garden path: The cognitive and biological basis for linguistic structures* (pp. 241–252). Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199677139.003.0012

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2016). The importance of linguistic typology for the neurobiology of language. *Linguistic Typology, 20*(3), 615–621. DOI: https://doi.org/10.1515/lingty-2016-0032

Butler, C. (2003). *Structure and function: Approaches to the simplex clause*. John Benjamins Publishing. DOI: https://doi.org/10.1075/slcs.63

Collart, A. (2024). Experimental linguistics embracing linguistic diversity: On the contributions of Formosan languages to models of sentence processing. In P. Li, E. Zeitoun, & R. De Busser (Eds.), *Handbook of Formosan languages: The indigenous languages of Taiwan, Part 2* (pp. 33–61). Brill. DOI: https://doi.org/10.1163/2772_5766_HFLO_COM_202230

Comrie, B. (1985). *Tense*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139165815

Dahl, Ö. (1990). Standard Average European as an exotic language. In J. Becher, G. Bernini, & C. Buridant (Eds.), *Toward a typology of European languages* (pp. 3–8). Mouton De Gruyter. DOI: https://doi.org/10.1515/9783110863178.3

Dahl, Ö. (2015, May 1–3). *How WEIRD are WALS languages?* [Conference presentation]. The Diversity Linguistics: Retrospect and Prospect conference. May 1-3, Leipzig, Germany.

Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511611896

Drummond, A. (2013). Ibex Farm. Access: https://farm.pcibex.net/

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer, & M. Haspelmath (Eds.), *WALS online* (v2020.3). DOI: https://doi.org/10.5281/zenodo.7385533

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance to cognitive science. *Behavioral and Brain Sciences, 32*, 429–492. DOI: https://doi.org/10.1017/S0140525X0999094X

Francis, E. J. (2022). *Gradient acceptability and linguistic theory*. Oxford University Press. DOI: https://doi.org/10.1093/oso/9780192898944.001.0001

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91*, 1357–1392. DOI: https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D. (2016). The neuroanatomical pathway model of language: Syntactic and semantic networks. In H. Gregory, & S. L. Small (Eds.), *Neurobiology of language* (pp. 349–356). Elsevier. DOI: https://doi.org/10.1016/B978-0-12-407794-2.00029-8

Givón, T. (2001). *Syntax: An introduction* (Vol. 1). John Benjamins Publishing. DOI: https://doi.org/10.1075/z.syn1

Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *NeuroImage, 20*(s1), 18–29. DOI: https://doi.org/10.1016/j.neuroimage.2003.09.013

Hagoort, P. (2016). MUC (Memory, Unification, Control): A model on the neurobiology of language beyond single word processing. In H. Gregory, & S. L. Small (Eds.), *Neurobiology of language* (pp. 339–347). Elsevier. DOI: https://doi.org/10.1016/B978-0-12-407794-2.00028-6

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2022). Glottolog 4.7. Max Planck Institute for Evolutionary Anthropology. DOI: https://doi.org/10.5281/zenodo.7398962 (Available online at http://glottolog.org, Accessed on 2023-05-26.)

Haspelmath, M. (2001). The European linguistic area: Standard Average European. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: An international handbook* (pp. 1492–1510). Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110194265-044

Hawkings, J. A. (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology, 25*(2), 87–107. DOI: https://doi.org/10.1016/j.newideapsych.2007.02.003

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*, 29. DOI: https://doi.org/10.1038/466029a

Jaeger, T. F., & Norcliffe, E. J. (2009). The cross-linguistic study of sentence production. *Language and Linguistics Compass, 3/4*, 866–887. DOI: https://doi.org/10.1111/j.1749-818X.2009.00147.x

Kandel, M., Yacovone, A., Slim, M., & Snedeker, J. (2022, March 24–27). *Webcams as windows to the mind: comparing web-based eye-tracking methods* [Conference presentation abstract]. The 35th Annual Conference on Human Sentence Processing, Santa Cruz, CA, United States.

Kidd, E., & Garcia, R. (2022). How diverse is child language acquisition? *First Language (Special issue: How diverse is child language acquisition research?), 42*(6), 703–735. DOI: https://doi.org/10.1177/01427237211066405

Kim, A. E., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205–225. DOI: https://doi.org/10.1016/j.jml.2004.10.002

Klein, W. (1994). *Time in language.* Routledge. DOI: https://doi.org/10.4324/9781315003801

Langlois, V. J., Ness, T., Kim, A. E., & Novick, J. M. (2023, March 9–11). *Using webcam eye-tracking to replicate subtle sentence processing effects* [Conference presentation abstract]. The 36th Annual Conference on Human Sentence Processing, Pittsburgh, PA, United States.

MacWhinney, B. (2022). The Competition Model: Past and future. In J. Gervain, G. Csibra, & K. Kovács (Eds.), *A life in cognition* (pp. 3–16). Springer. DOI: https://doi.org/10.1007/978-3-030-66175-5_1

Majid, A., & Levinson, S. C. (2010). WEIRD languages have misled us, too. *Behavioral and Brain Sciences, 33*(2/3), 103. DOI: https://doi.org/10.1017/S0140525X1000018X

Malik-Moraleda, S., Ayyash, D., Gallée, J., Affourtit, J., Hoffmann, M., Mineroff, Z., Jouravlev, O., & Fedorenko, E. (2022). An investigation across 45 languages and 12 language families reveals a universal language network. *Nature Neuroscience, 25*, 1014–1019. DOI: https://doi.org/10.1038/s41593-022-01114-5

Mazerolle, M. J. (2020) AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.3-1. https://cran.r-project.org/package=AICcmodavg.

Meyer, R. (2023). *Iranian syntax in Classical Armenian: The Armenian perfect and other cases of pattern replication.* Oxford Academic. DOI: https://doi.org/10.1093/oso/9780198851097.001.0001

Musan, R. (1997). Tense, predicates, and lifetime effects. *Natural Language Semantics, 5*(3), 271–303. DOI: https://doi.org/10.1023/A:1008281017969

Norcliffe, E. J., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience, 30*(9), 1009–1032. DOI: https://doi.org/10.1080/23273798.2015.1080373

Ono, H., Kim, J., Sato, M., Tang, A. A.-Y., & Koizumi, M. (2020). Syntax and processing in Seediq: A behavioral study. *Journal of East Asian Linguistics, 29*(2), 237–258. DOI: https://doi.org/10.1007/s10831-020-09207-7

R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Sato, M., Niikuni, K., Schafer, A. J., & Koizumi, M. (2020). Agentive versus non-agentive motions immediately influence event apprehension and description: An eye-tracking study in a VOS language. *Journal of East Asian Linguistics, 29*(2), 211–236. DOI: https://doi.org/10.1007/s10831-020-09205-9

Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2009). When semantic P600s turn into N400s: On cross-linguistic differences in online verb-argument linking. In M. Horne, M. Lindgren, M. Roll, K. Alter, & J. v. K. Torkildsen (Eds.), *Brain talk: Discourse with and in the brain. Papers from the first Birgit Rausing Language Program Conference in Linguistics* (pp. 75–97). Birgit Rausing Language Program.

Schwarz, F., & Zehr, F. (2021). Tutorial: Introduction to PCIbex – An open-science platform for online experiments: Design, data-collection and code-sharing. *Proceedings of the Annual Meeting of the Cognitive Science Society, 43*.

Speed, L. J., Wnuk, E., & Majid, A. (2018). Studying psycholinguistics out of the lab. In A. M. B. de Groot, & P. Hagoort (Eds.), *Research methods in psycholinguistics and the neurobiology of language: A practical guide* (pp. 190–207). Blackwell. DOI: https://doi.org/10.1002/9781394259762.ch10

Van Valin, R. (2005). *Exploring the syntax-semantics interface*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511610578

Wagers, M., & Chung, S. (2023). Language processing experiments in the field. In J. Sprouse (Ed.), *The Oxford handbook of experimental syntax* (pp. 491–512). Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198797722.013.15

Whalen, D. H., & McDonough, J. (2015). Taking the laboratory into the field. *Annual Review of Linguistics*, *1*, 395–415. DOI: https://doi.org/10.1146/annurev-linguist-030514-124915

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. DOI: https://doi.org/10.1007/978-0-387-98141-3

Yano, M., Niikuni, K., Ono, H., Sato, M., Tang, A. A.-Y., & Koizumi, M. (2019). Syntax and processing in Seediq: An event-related potential study. *Journal of East Asian Linguistics*, *28*(4), 395–419. DOI: https://doi.org/10.1007/s10831-019-09200-9

Zeitoun, E., Huang, L. M., Yeh, M. M., Chang, A. H., & Wu, J.-l. J. (1996). The temporal, aspectual, and modal systems of some Formosan languages: A typological perspective. *Oceanic Linguistics*, *35*(1), 21–56. DOI: https://doi.org/10.2307/3623029