



A bibliometric comparison of grammatical development in spoken and sign languages: Thematic evolution and persistent keywords^{*}

Sunghwa Lee^{**a} · Myoungho Ha^b · Byong-Rae Ryu^c · Se-Eun Jhang^{***a}
(National Korea Maritime & Ocean University^a · Silla University^b ·
Chungnam National University^c)

Lee, Sunghwa, Myoungho Ha, Byong-Rae Ryu, and Se-Eun Jhang. 2025. A bibliometric comparison of grammatical development in spoken and sign languages: Thematic evolution and persistent keywords. *Linguistic Research* 42(Special Edition): 109-147. This study offers a bibliometric comparison of grammatical research in spoken and sign languages, focusing on thematic evolution and key anchors. Drawing on 22,947 spoken and 712 sign language articles (1989–2024), we applied Biblioshiny to extract and visualize bibliometric indicators. Research periods were segmented using three data-driven metrics—Annual Scientific Production, Average Citations per Year, and Trend Topics—enabling modality-specific periodization. Thematic evolution analyses revealed a parallel three-phase trajectory in both modalities: foundational, growth, and expansion. Spoken language research showed greater diversification, especially in psycholinguistic and syntactic domains, whereas sign language studies remained centered on modality-specific areas such as visual processing and accessibility. Persistent keyword analysis underscored shared anchors (e.g., acquisition, syntax, perception) alongside modality-specific emphases. Despite dataset limits, this study advances a multimodal framework for grammar research, highlights gaps in sign language syntax and psycholinguistic modeling, and outlines directions for cross-modal integration. (National Korea Maritime & Ocean University · Silla University · Chungnam National University)

* This research was supported by the 2025 Digital Humanities Utmost Sharing System (D-HUSS, NRF-2023-1479-01) at Chungnam National University. An early version of this paper was presented orally at the 3rd Workshop on Interdisciplinary Research on the Acquisition of Korean Sign Language at Chungnam National University on 27 June 2025. We are grateful to the two anonymous reviewers for their incisive comments and suggestions, which substantially improved the quality of this manuscript.

** First author

*** Corresponding author

Keywords Bibliometric analysis, thematic evolution, sign languages, grammatical development, author keywords

1. Introduction

Languages, whether conveyed through sound or gesture, are governed by complex grammatical structures that enable efficient and meaningful communication. Spoken languages, transmitted through auditory-vocal channels, and sign languages, expressed through visual-gestural modalities, differ fundamentally in their modes of expression. However, both systems share core linguistic properties, including hierarchical syntax, morphological structure, and rules of agreement (Sandler and Lillo-Martin 2006). These shared properties have motivated cross-modal comparisons and raised important theoretical questions regarding the universality and variability of grammar across language types.

In recent decades, spoken language research has benefited from extensive theoretical foundations, including generative grammar (Chomsky 1965), constraint-based approaches (Bresnan 2001), and developments in psycholinguistics and neurolinguistics. In contrast, sign language linguistics, while advancing significantly, remains underrepresented in large-scale empirical studies and bibliometric meta-analyses. Much of the existing literature on sign language grammar has focused on language acquisition and syntactic agreement (Lillo-Martin and Meier 2011), but its integration with computational methods, neurocognitive approaches, and large corpora is still emerging.

Bibliometric analysis offers a powerful lens for tracing the evolution of research topics over time. Specifically, thematic evolution analysis enables the identification of persistent research themes, emergent topics, and shifts in scholarly focus. However, one notable limitation in prior bibliometric studies is the often arbitrary division of thematic periods—typically based on round-number decades or uniform intervals—rather than data-driven trends (Parlina et al. 2020; Crosthwaite et al. 2023; Hajar and Karakus 2025). For example, several studies in linguistics and applied domains have segmented timespans by fixed decades (e.g., 2000–2009, 2010–2019) without reference to actual inflection points in publication or citation patterns. This study addresses that gap by introducing a more rational and empirically grounded approach

to period segmentation. Specifically, it defines the research periods based on three complementary bibliometric indicators—Annual Scientific Production, Average Citations per Year, and Trend Topics—thus establishing them as meaningful and evidence-based analytical units (Liu et al. 2024; Liu and Jhang 2024; Lee and Jhang 2025; Liu 2025).

Moreover, this study contributes beyond quantitative trends by offering an in-depth thematic interpretation through period-specific thematic maps. While many bibliometric studies emphasize publication and citation counts or keyword co-occurrence networks, they often stop short of interpreting thematic evolution in detail (Zupic and Čater 2015; Donthu et al. 2021). By contrast, this study uses thematic evolution not only to segment the periods empirically, but also to trace the conceptual maturation of grammar research within and across modalities.

Moreover, this study presents a novel approach by examining thematic maps for each identified period and comparing persistent themes across the development of spoken and sign language grammar. Through this method, it analyzes the grammatical research trajectories of spoken and sign languages in parallel, offering insights into modality-specific developments and shared linguistic concerns. To the best of our knowledge, no prior research has undertaken a comparative thematic evolution analysis that systematically explores both persistent keywords and period-specific thematic structures across spoken and sign language corpora. This dual-level exploration enables a deeper understanding of how grammatical inquiry evolves differently depending on language modality and technological or theoretical shifts.

This research thus seeks to uncover developmental trajectories, identify research gaps, and propose grounded directions for future inquiry in sign language linguistics, contributing not only to our understanding of grammatical evolution but also to the broader theoretical discussion on Universal Grammar.

To achieve these objectives, the study is guided by the following research questions, each of which corresponds to a core analytical aim of the thematic evolution framework applied in this work:

- Research Question 1: How have the key grammatical themes evolved differently in spoken and sign languages between 1989 and 2024, as revealed through thematic evolution analysis?
- Research Question 2: What persistent keywords appear consistently across all three periods (P1–P3) in both spoken and sign languages, and what do they

imply about universal grammatical structures?

- Research Question 3: Based on the results of the thematic evolution analysis, what are the underexplored areas and priority methodological directions that should be strengthened in sign language grammar research?

2. Literature reviews

2.1 Bibliometric studies on grammar in spoken and sign languages

Bibliometric analysis has become an important method for identifying research trends in linguistics. However, studies that focus specifically on grammar—especially comparing spoken and sign languages—are still rare. A few prior studies, though not directly centered on grammar, provide valuable methodological foundations. For instance, Crosthwaite et al. (2023) conducted a bibliometric analysis of corpus linguistics using Scopus-indexed data from the past two decades. Their application of keyword co-occurrence and thematic mapping provides useful procedures that can be adapted for grammar-focused research.

In the field of sign language, research is even more limited. Narenthiran et al. (2016) analyzed sign language publications between 1991 and 2015, but their focus was primarily descriptive, and they did not isolate grammar as a central unit of analysis. More recently, Lee and Jhang (2025) presented thematic insights into sign-language research at a Korean linguistics conference, again without making grammar the comparative pivot. These works highlight both the value of bibliometric methods and the persistent gap: no study has systematically examined grammatical development across spoken and sign languages in a comparative bibliometric framework.

This gap is important because grammar is where linguistic universals and modality-specific realizations most clearly intersect. Spoken languages primarily rely on linear, auditory-prosodic cues, whereas sign languages exploit visual-spatial mechanisms such as simultaneity and non-manual markers (Sandler and Lillo-Martin 2006; Lillo-Martin and Meier 2011). A bibliometric approach that places grammar at the center therefore offers a principled way to track how each modality has developed thematically and how research priorities diverge or converge over time.

2.2 Methodological issues in bibliometric reviews

Bibliometric reviews, while powerful, also have methodological limitations. Previous studies often rely heavily on raw publication and citation counts, which risks overstating quantitative growth while obscuring conceptual change (Zupic and Čater 2015; Donthu et al. 2021). In addition, reliance on single databases (e.g., Web of Science or Scopus) may produce coverage bias (Arik 2014), and inconsistencies in keyword indexing can distort thematic maps (Parlina et al. 2020). Temporal segmentation also presents challenges: dividing periods arbitrarily by decade may mask more meaningful turning points in research trajectories.

To address these concerns, our study adopts bibliometrix/Biblioshiny (Aria and Cuccurullo 2017), which provides transparent science-mapping functions such as co-word analysis, thematic evolution, and clustering with reproducible parameter settings. For data preprocessing, we employed Python (spaCy; Honnibal et al. 2020) for lemmatization and developed custom dictionaries to unify sign-language variants and resolve inconsistencies (e.g., *nonmanual/non-manual*, *ASL/American Sign Language*). Rule-based normalization of acronyms, hyphenation, and frequent spelling variants was also applied. These steps ensure that grammatical terminology is consistently represented, reducing noise in cluster formation and persistent keyword detection.

Another methodological concern is the imbalance between spoken- and sign-language datasets. Bibliometric practice offers several normalization strategies to mitigate this: fractional counting to equalize contributions of multi-authored records (Perianes-Rodríguez et al. 2016), field-normalization to account for disciplinary differences (Waltman 2016), and association-strength normalization to stabilize co-occurrence networks (Van Eck and Waltman 2010). We acknowledge these standard techniques but deliberately chose not to apply them, because our central aim is to capture the empirical disparity itself rather than to impose an artificial balance. This choice allows us to foreground structural differences in productivity and thematic diversification between the two modalities, though we also mark it as a limitation and discuss it further in Section 5.4.

Finally, we note that quantity alone does not guarantee conceptual change. Accordingly, our analysis interprets clusters and persistent keywords against theoretical

constructs of grammar (e.g., *agreement*, *inflection*, *prosody*) and methodological imports from psycholinguistics and AI. This combined approach enables us to identify not only growth but also qualitative shifts—such as the expansion of psycholinguistic constructs in spoken-language grammar research and the technology-driven themes (*recognition*, *machine translation*, *avatar*) emerging in sign-language grammar studies.

3. Data and methodology

3.1 Data collection

Bibliographic records were retrieved from the Web of Science (WoS) Core Collection on June 22, 2025, to enable a comparative analysis of grammatical research in spoken and sign languages. To maximize coverage while preserving conceptual specificity, we used Boolean operators, wildcard truncation, and field-tagged topic searches (TS=) across titles, abstracts, and keywords.

For sign language grammar, the query targeted studies that contained both sign-language terms and grammatical terminology in the Topic field: TS=("sign language*" OR "signed language*") AND TS=(grammar* OR synt* OR morphosynt* OR morpholog*). Wildcard operators (e.g., *grammar*, *synt*, *morphosynt*, *morpholog*) were applied to capture inflectional and derivational variants (e.g., *grammar/grammatical*, *syntax/syntactic*, *morphosyntax/morphosyntactic*, *morphology/morphological*). To reflect the interdisciplinary nature of sign language research while maintaining linguistic relevance, four WoS categories were included: Linguistics, Language & Linguistics, Computer Science—Artificial Intelligence, and Computer Science—Interdisciplinary Applications.

For spoken language grammar, a parallel query was applied while explicitly excluding sign language to ensure dataset independence: TS=(grammar* OR synt* OR morphosynt* OR morpholog*) NOT TS=("sign language*" OR "signed language*"). The dataset was restricted to the Linguistics category to maintain a disciplinary focus.

For both modalities, only journal articles were included, and the search was limited to English-language publications. The analysis covered 1989–2024 for sign language and 1990–2024 for spoken language, reflecting the earliest available years in WoS for each domain. This procedure yielded 22,947 spoken language articles and 712

sign language articles, which served as the basis for subsequent preprocessing and analysis.

3.2 Data preprocessing

All retrieved records were exported from the Web of Science in plain-text format and converted to Excel for preprocessing. A Python-based pipeline was implemented to ensure consistency, lexical normalization, and reproducibility of the keyword data prior to bibliometric analysis. Preprocessing was conducted in Python 3.11 using the libraries pandas (v2.2), spaCy (v3.7), and NLTK (v3.8), which supported tokenization, lemmatization, and keyword standardization.

A rule-based lemmatizer from spaCy was applied to reduce inflected or derived forms to their base lemma. For example, phonetic and phonetics were reduced to phonetics, and signed language was standardized to sign language. The spaCy (v3.7) pipeline also assigned part-of-speech tags to all tokens, including NOUN, VERB, ADJ, ADV, PROPN, DET, and CCONJ, which guided the lemmatization process and ensured consistent handling of grammatical categories. Beyond lemmatization, a set of unification rules was applied to handle orthographic, morphological, and terminological variants. Most inflectional and orthographic variants (e.g., grammar/grammatical, sign-language/sign languages) were successfully normalized.

Representative examples of these rules are presented in Table 1. The rules addressed plural/singular alternations, spelling variations, and synonymous expressions, resulting in 18 standardized mappings in total. This procedure ensured that equivalent terms were treated consistently across the dataset, thereby minimizing fragmentation in co-word networks.

Table 1. Examples of lexical normalization rules applied in preprocessing

Original form	Standardized form
phonetic	phonetics
phonetics	phonetics
grammatical	grammar
grammars	grammar
syntactic	syntax
morphosyntactic	morphosyntax
morphological	morphology

sign languages	sign language
sign-language	sign language
signed language	sign language
L2 acquisition	second language acquisition
children	child
deaf children	deaf-children
facial expressions	facial expression
non-manual markers	nonmanual
agreement verbs	agreement
sentence processing	comprehension
machine translations	machine translation

While most variants were normalized automatically, abbreviations such as ASL versus american sign language were not consistently unified in the dataset. These cases required explicit post-processing, and in the present analysis they remained partially unresolved.

Finally, all keyword entries were lowercased, punctuation and diacritics were removed, and duplicates were eliminated to create a standardized dataset. It is important to distinguish between lexical normalization, which was applied as described above, and dataset-size normalization, which was deliberately not implemented. The issue of corpus imbalance between spoken and sign language studies is addressed in Section 5.4.

3.3 Biblioshiny analysis

Bibliometric analyses were conducted using Biblioshiny (v5.0), the web-based interface of the bibliometrix R package (v5.0) (Aria and Cuccurullo 2017). All procedures were implemented in R 4.5.0 (2025-04-11, ucrt), ensuring compatibility with the latest release environment. Biblioshiny was selected for its standardized implementation of co-word analysis, thematic mapping, and thematic evolution, which allowed for transparent and replicable procedures across both spoken and sign language datasets.

The analyses were conducted in Biblioshiny¹ using the All Keywords (KW_Merged) field, which integrates Author Keywords and Keywords Plus. Co-word networks were

¹ While representative screenshots of the Biblioshiny interface were not reproduced here due to space limitations, the analyses strictly followed the built-in modules (e.g., Annual Scientific Production, Average Citations per Year, Trend Topics). The corresponding outputs are documented in Section 4.

generated using the Walktrap clustering algorithm, with association strength specified as the coupling measure. Keyword co-occurrence matrices were normalized using Salton’s cosine index, a widely adopted similarity coefficient in bibliometric research.

Thematic map clusters were constructed with a minimum cluster frequency of five, one-word n-grams, and a cluster size parameter of 0.3. For each cluster, centrality and density were calculated using standard graph-theoretic measures, namely betweenness centrality, closeness centrality, and PageRank centrality. These indices provided complementary perspectives on the structural importance and internal cohesion of thematic clusters across the three research periods.

3.4 Dataset overview

The final dataset consisted of two corpora representing grammatical research in spoken and sign languages. Their characteristics are summarized in Table 2.

Table 2. Summary of dataset characteristics for spoken and sign language grammar research²

Metric	Spoken Languages	Sign Languages
Timespan	1990-2024	1989-2024
Number of Documents	22,947	712
Average Citations per Document	22.4	17.95
Annual Growth Rate (%)	8.66%	11.76%
Keywords Plus (n)	9,370	606
Author’s Keywords (n)	31,767	1,566
Average Authors per Document	2.07	2.86
International Collaboration (%)	19.50%	19.94%

As shown in Table 2, spoken language grammar research is far larger in scale, but sign language research has exhibited a higher annual growth rate and a comparable level of international collaboration. The keyword coverage also illustrates this imbalance: while spoken language studies include far more Keywords Plus and Author’s Keywords, the average number of authors per paper is slightly higher in sign language studies, suggesting stronger collaborative practices within a smaller community.

² All values are based on bibliographic data retrieved from the Web of Science Core Collection and analyzed using Biblioshiny.

Taken together, these metrics underscore the structural imbalance between the two domains. Although research into the grammar of spoken languages has a longer tradition, larger data pools and broader institutional support, research into the grammar of sign languages demonstrates greater growth momentum and distinctive patterns of collaboration, despite being smaller in scale. The methodological implications of this imbalance are addressed in Section 5.4.

4. Results and discussions

4.1 Period segmentation overview

To support the thematic evolution analysis, each dataset was divided into three periods using three bibliometric indicators: Annual Scientific Production, Average Citations per Year, and Trend Topics. While publication and citation trends revealed overall growth and impact, Trend Topics played a key role in identifying thematic shifts. Notably, the emergence of new keywords in specific years (e.g., 2012 for spoken language, 2016 for sign language) guided the period boundaries. Subsections 4.1.1 and 4.1.2 explain the segmentation process for each modality.

4.1.1 Spoken languages: Period division

Publication output in spoken language grammar research steadily increased from the early 1990s and surged through the 2000s and 2010s, with consistently high levels into the 2020s. As shown in Figure 1, annual production expanded from fewer than 200 articles in the early 1990s to over 1,200 articles by the early 2020s. Citation impact peaked around 2003, with a secondary rise in the early 2000s, before gradually declining in recent years as the publication base expanded.

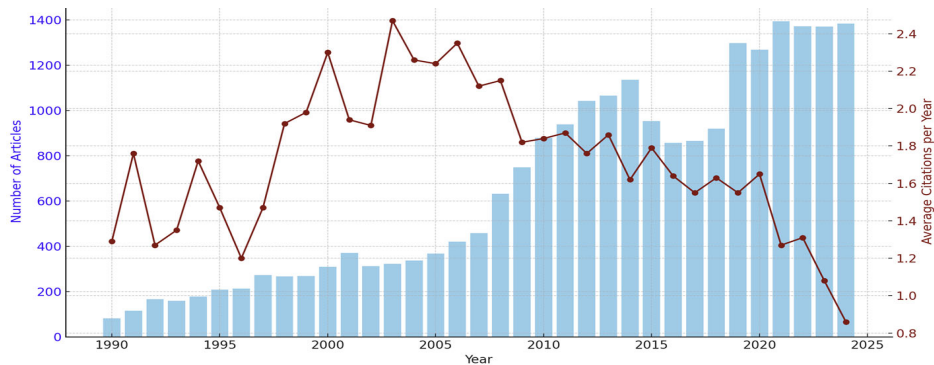


Figure 1. Annual scientific production and average citations per year for spoken language grammar studies (1990–2024)³

The segmentation of research periods was guided by a combination of three criteria: (i) publication productivity, (ii) citation trends, and (iii) the emergence of new research topics. While the visualization of trend topics is provided in Figure 2 (See Appendix Figure 2 for readability of topic items), here we explicitly describe the new topics that signaled thematic shifts.

- Around 2000, new research directions emerged that connected grammatical inquiry to clinical and developmental perspectives as well as detailed syntactic processing. Topics included *disorder*, *anomia*, *nonstutters*, *syntactic comprehension*, *syntactic ambiguity resolution*, *phrase structure*, and *cognitive-development*. These keywords indicate the field's turn toward psycholinguistic processing, clinical populations, and theoretical syntax.
- In 2012, another shift occurred, this time consolidating core domains of linguistic theory and broadening comparative perspectives. Newly introduced topics were *spanish*, *knowledge*, *semantics*, and *agreement*, reflecting a strengthened focus on semantics, morphosyntactic relations, and cross-linguistic inquiry.

Based on the combined evidence of productivity, citation impact, and emerging themes, the following three periods were established:

³ Blue bars represent the number of publications per year; the brown line represents average citations per document per year.

- Period 1 (1989–1999): A formative phase characterized by low research output and a foundational focus on syntax and language acquisition, laying the groundwork for subsequent empirical expansion.
- Period 2 (2000–2011): A consolidation phase marked by the integration of psycholinguistic processing, clinical perspectives, and syntactic theory, reflecting a shift toward experimental and interdisciplinary approaches.
- Period 3 (2012–2024): A phase of semantic and cross-linguistic broadening, with increased attention to meaning, morphosyntactic relations, and specific language studies, indicating a diversification of grammatical inquiry.

4.1.2 Sign languages: Period division

In sign language grammar research, publication output remained modest until the early 2000s but began to rise steadily from 2003 onward. As shown in Figure 3, annual production increased gradually, while citation impact fluctuated, with an early peak in 1998 and another noticeable rise around 2017–2018.

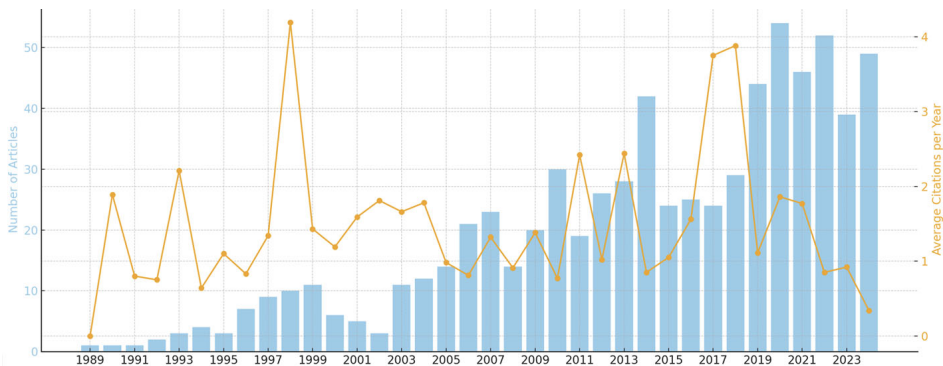


Figure 3. Annual scientific production and average citations per year for sign language grammar studies

The segmentation of research periods was informed by the combined evidence of productivity, citation patterns, and the introduction of new topics. While the visualization is provided in Appendix Figure 4 for readability of topic items, the thematic shifts can be summarized as follows.

- In 2004, sign language studies expanded their scope by focusing on *signed language*, *acquisition*, and *American Sign Language*, signaling a stronger orientation toward language learning and the detailed investigation of specific sign languages.
- In 2016, new directions emerged that emphasized *age*, *iconicity*, and *machine translation*. These topics highlighted the integration of developmental and cognitive perspectives, as well as the increasing application of computational approaches to accessibility and translation.

On the basis of these combined developments, three phases of research can be identified:

- Period 1 (1989–2003): A formative phase with modest research output, focusing on the foundational description of sign language structure and early studies of acquisition.
- Period 2 (2004–2015): A consolidation phase marked by steady growth and increasing attention to sign language acquisition and in-depth investigation of individual sign languages, establishing a stronger empirical base for grammatical studies.
- Period 3 (2016–2024): A diversification phase characterized by the incorporation of developmental and cognitive perspectives alongside technological applications, reflecting a broader interdisciplinary and application-oriented orientation in sign language grammar research.

4.1.3 Comparative summary of periodization between modalities

Despite differing developmental contexts, both spoken and sign language grammar research exhibit a shared three-phase trajectory: formation, growth, and expansion. Early studies in both modalities emphasized core theoretical constructs such as syntax and acquisition. These were followed by an experimental turn, marked by rising interest in psycholinguistic themes like memory, modality, and comprehension. In the most recent period, attention has shifted toward interdisciplinary engagement, particularly with artificial intelligence and cognitive neuroscience, as reflected in keywords such as machine translation, recognition, and avatar.

Yet important differences remain. Spoken language grammar research, grounded in a longer academic tradition and a substantially larger dataset (22,947 articles), allowed for clear segmentation based on bibliometric indicators—citation peaks (e.g., 2003) and bursts of new keywords (e.g., around 2012)—as shown in modules like Annual Scientific Production and Trend Topics. While quantitative indicators such as Annual Scientific Production and Average Citations per Year revealed overall growth, their irregular patterns did not yield clear breakpoints. Instead, qualitative interpretation was based on the onset of new themes in trend topics, where the appearance of modality and iconicity in the mid-2000s, and later the rise of technology-related terms such as machine translation, recognition, and avatar after 2016, served as decisive markers of thematic transitions. These thematic onsets provided more meaningful turning points than numerical thresholds, guiding the division into exploratory, consolidative, and applied phases.

Furthermore, key thematic transitions in sign language research consistently lag behind those in spoken language. Experimental and psycholinguistic topics appeared in spoken language studies around 2000 but only gained traction in sign language by the mid-2000s. Similarly, technology-driven themes emerged in spoken language after 2012 but not until around 2016 in sign language.

This temporal lag underscores enduring structural disparities—limited institutional support, smaller data pools, and delayed adoption of interdisciplinary methods—that continue to constrain the development of sign language grammar research. Bridging this gap is not merely an academic concern; it is essential for ensuring theoretical equity and advancing a more inclusive linguistic science.

4.2 Thematic evolution of grammatical research

To examine how core grammatical research themes have developed over time, this study performed a thematic evolution analysis using the All Keywords field. By generating thematic maps for each period and identifying overlapping keywords across adjacent periods, the analysis reveals dominant and emerging themes as well as their transitions. These results illustrate conceptual maturation, diversification, and interdisciplinary convergence within each modality.

4.2.1 Thematic evolution in spoken language grammar research

4.2.1.1 Period 1 (1989–1999)

Thematic analysis of the first period reveals a prominent motor theme consisting of language, syntax, and acquisition, located in the upper-right quadrant of the thematic map (Figure 5). This indicates that foundational grammatical constructs formed a well-developed and central cluster in early spoken language research. These themes reflect the field’s initial theoretical consolidation, aligning with the formative stage identified in Section 4.1.1.

In addition to this dominant cluster, other topics such as *comprehension*, *lexical access*, *aphasia*, and *second language acquisition* appeared in adjacent quadrants. For example, comprehension and discourse appeared as basic themes (lower-right quadrant), while *lexical access* and *retrieval* were positioned in the emerging cluster zone (lower-left quadrant). This suggests that while core syntactic and acquisition-focused themes anchored the period, psycholinguistic and cognitive lines of inquiry were beginning to surface at the margins.

Overall, the thematic structure of this period was relatively cohesive, with theoretical grammar playing a central role and peripheral topics emerging gradually.

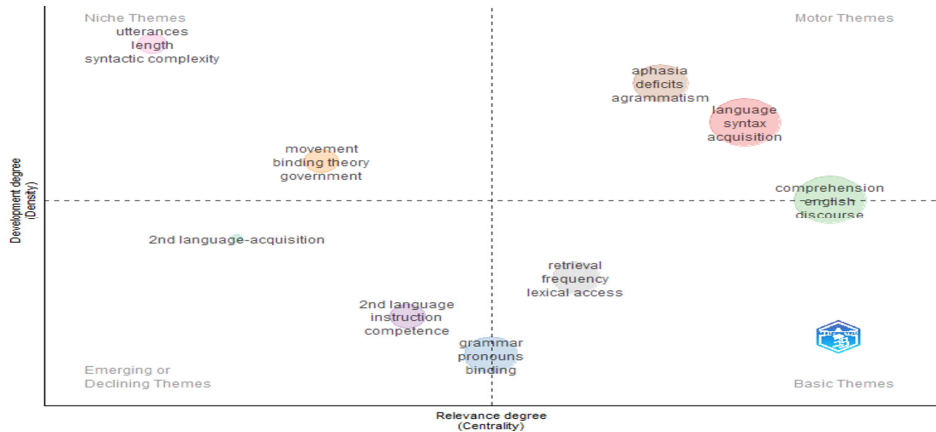


Figure 5. Thematic map of period 1 (1989–1999) for spoken language grammar⁴

4 Thematic map quadrants represent topic importance (centrality) and development (density). Motor themes (upper-right) are well-developed and central; upper-left themes are developed but disconnected; lower-left themes are emerging or fading; lower-right themes are basic but broadly relevant.

4.2.1.2 Period 2 (2000–2011)

Period 2 maintained a strong emphasis on core syntactic research, with limited thematic diversification. As shown in Figure 6, the thematic structure was dominated by a single large Motor Theme cluster consisting of *language*, *syntax*, and *comprehension*, reflecting the field's continued focus on sentence-level processing and syntactic representation.

Other keywords such as *acquisition*, *grammar*, and *English* appeared in the Emerging/Declining quadrant, indicating that while still relevant, these topics had a lower density and centrality within the research network. This pattern suggests that the field remained structurally cohesive but did not undergo significant thematic branching during this period. The lack of distinct secondary clusters implies that new subfields or interdisciplinary links had yet to gain prominence.



Figure 6. Thematic map of period 2 (2000–2011) for spoken language grammar

4.2.1.3 Period 3 (2012–2024)

In Period 3, the thematic structure of spoken language grammar research remained centered on its core cluster. As seen in Figure 7 below, the terms *language*, *syntax*, and *comprehension* continued to occupy the Motor Themes quadrant, underscoring the persistent centrality of traditional syntactic concerns. Although terms such as *acquisition*, *grammar*, and *English* remained present in the Emerging/Declining quadrant, they did not show increased density or connectivity, suggesting that these

areas did not develop into major thematic hubs. Contrary to expectations of diversification, the field in this period appears to have reconsolidated around established core themes. This indicates that spoken language grammar research continued to prioritize stability and depth in central topics rather than expanding into novel or peripheral areas.

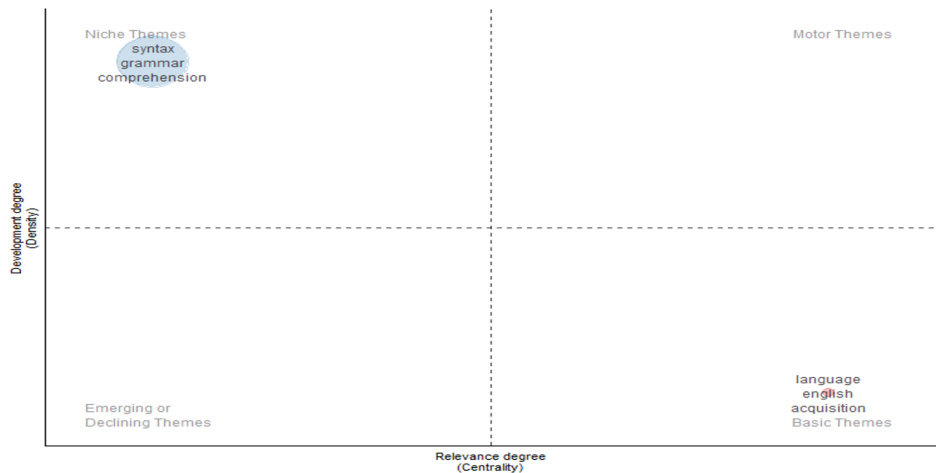


Figure 7. Thematic map of period 3 (2012–2024) for spoken language grammar

4.2.1.4 Overall thematic landscape

An examination of the overall thematic map for spoken language grammar research across all three periods reveals a relatively stable and centralized thematic structure (see Figure 8 below). The cluster of *language*, *English*, and *acquisition* remains a dominant motor theme, maintaining both high centrality and density. This consistency indicates a core research focus that has persisted throughout the field’s development. In contrast, topics such as *instruction*, *syntactic complexity*, and *learners* appear in the lower-left quadrant, reflecting low development and marginal influence across the dataset. The thematic landscape suggests that while the field has diversified within certain subdomains over time, its central research agenda has remained largely conservative. This continuity contrasts with the more dynamic shifts observed in sign language research, discussed in the following section.

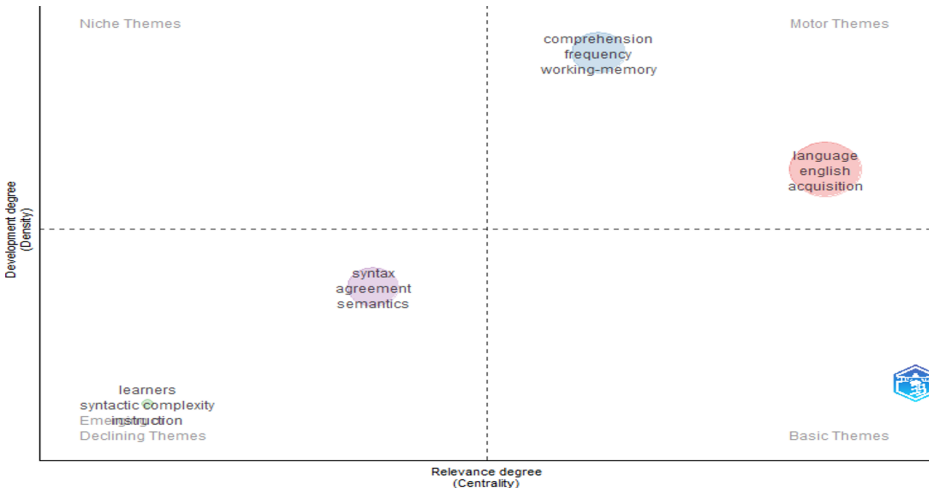


Figure 8. Thematic map of spoken language grammar across all periods

Despite the diversification observed in peripheral subfields, spoken language grammar research has maintained a relatively stable core of themes—most notably *language*, *acquisition*, and *English*—throughout its evolution. This enduring centrality suggests a well-established disciplinary foundation and a cumulative research trajectory. Thematic developments in sign language grammar research, however, reveal a different pattern, as discussed in the next section.

4.2.2 Thematic evolution in sign language grammar research

4.2.2.1 Period 1 (1989–2003)

During the first period of sign language grammar research, the thematic landscape was centered on a small but coherent set of highly relevant topics. As shown in Figure 9 below, a dominant motor theme cluster emerged, consisting of keywords such as *sign language*, *acquisition*, *American Sign Language*, *language*, *speech*, and *facial expression*. These terms represent the field's early focus on foundational aspects of language structure and acquisition within the sign modality, often in relation to spoken language comparisons.

Adjacent to this cluster in the motor quadrant are several cognitively-oriented keywords including *memory*, *recognition*, *representation*, *plasticity*, *bilinguals*, and

words. Their position within the same quadrant, though slightly more peripheral, reflects the growing incorporation of neurocognitive perspectives into grammatical research even at this early stage.

In contrast, terms such as *agreement*, *rhythm*, *animation*, and *American Sign Language (ASL)* appear in the lower-left quadrant as emerging or declining themes, indicating low centrality and developmental maturity. Additionally, *hearing* was positioned in the lower-right basic themes quadrant, suggesting its peripheral relevance but frequent citation. Lastly, *hidden Markov models* was the only theme to appear near the niche quadrant, hinting at early experimental approaches to modeling sign language structure, though not yet widely integrated.

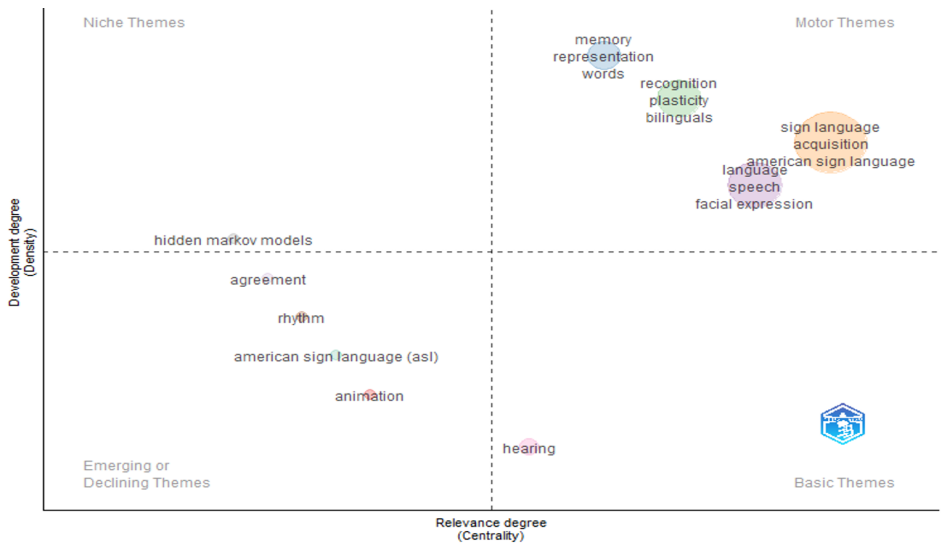


Figure 9. Thematic map of period 1 (1989–2003) for sign language grammar

4.2.2.2 Period 2 (2004–2015)

In the second period, sign language grammar research experienced both expansion and thematic dispersion. As shown in Figure 8, the thematic map reveals a dense and fragmented landscape, with keywords scattered across all four quadrants. Unlike the cohesive structure of Period 1, Period 2 presents challenges in interpretation due to the overlapping positions of numerous keywords and the absence of clearly defined thematic clusters.

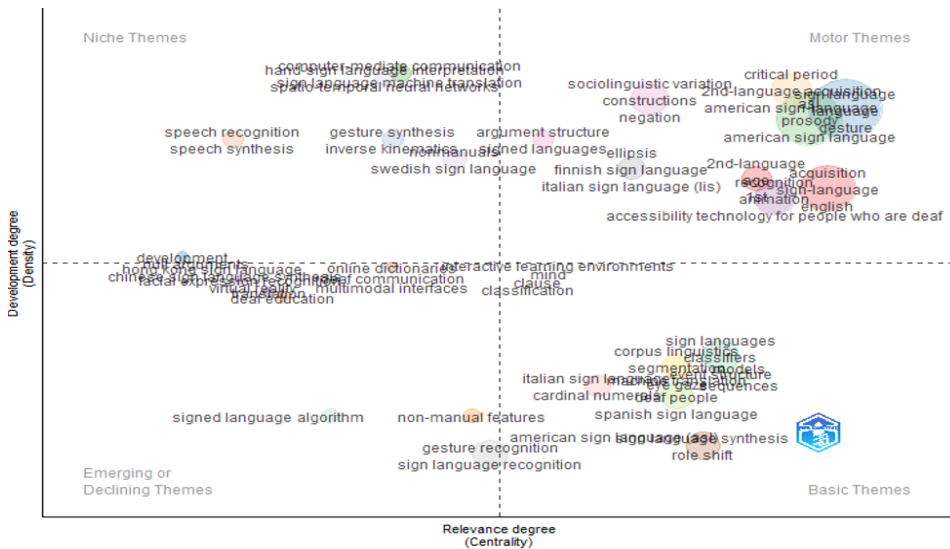


Figure 10. Thematic map of period 2 (2004–2015) for sign language grammar

Figure 10 above shows that a core motor theme remains present in the upper-right quadrant, composed of *sign language*, *language*, *gesture*, *acquisition*, and *American Sign Language*, suggesting the persistence of foundational topics. However, many additional keywords—such as *prosody*, *second-language acquisition*, *facial expression*, *negation*, *multimodal interfaces*, and *deaf education*—are widely dispersed across the map, reflecting a proliferation of emerging interests that had not yet coalesced into cohesive research strands.

To supplement this complexity, a frequency-based keyword analysis was conducted (Table 3 below), highlighting the dominant terms that structured the thematic field during this period. While *sign language* and *language* remained the most frequently occurring keywords, other terms such as *gesture*, *ASL*, and *prosody* also appeared prominently, indicating growing attention to multimodal features and expressive aspects of sign language. Beyond frequency, centrality measures highlight how each keyword functions within the research network. For instance, *gesture* showed high betweenness, indicating its role as a connector between different thematic areas, while *prosody* scored high on closeness, reflecting its conceptual proximity to many other core topics.

Table 3. Top 15 most frequent keywords in sign language grammar research (2004–2015)⁵

Rank	Keywords	Occurrences	Cluster	Betweenness	Closeness	PageRank
1	sign language	51	1	0.338	0.607	0.089
2	language	22	1	0.157	0.572	0.053
3	ASL	19	1	0.081	0.521	0.04
4	acquisition	15	1	0.065	0.495	0.037
5	gesture	14	2	0.062	0.511	0.034
6	prosody	13	2	0.054	0.476	0.032
7	deaf	12	3	0.049	0.460	0.029
8	speech	11	2	0.045	0.447	0.028
9	facial expression	10	1	0.043	0.438	0.026
10	bimodal	9	3	0.038	0.426	0.025
11	critical period	9	1	0.037	0.421	0.024
12	deaf education	8	3	0.035	0.418	0.023
13	negation	8	2	0.033	0.412	0.022
14	mouthings	7	2	0.030	0.407	0.021
15	nonmanual	7	2	0.029	0.404	0.021

This thematic dispersion can be attributed to two main factors. First, the increasing interdisciplinarity of the field brought in diverse perspectives from linguistics, cognitive science, education, and computer science, resulting in a broadened yet less integrated research landscape. Second, although the volume of research grew compared to Period 1, the overall dataset was still relatively small, limiting the stability of cluster formation in co-word analysis.

Overall, Period 2 represents a transitional phase characterized by the coexistence of core grammatical themes and an expanding array of exploratory directions, setting the stage for the more consolidated shifts that follow in Period 3.

4.2.2.3 Period 3 (2016–2024)

In the third period, sign language grammar research underwent a notable shift toward applied and technology-driven directions. As illustrated in Figure 11 below, the thematic map reveals a central motor cluster built around *sign language*, *language*, *recognition*, *translation*, *avatar*, *machine translation*, and *accessibility*. These terms

⁵ Occurrences indicate how frequently each keyword appeared during the period. Cluster refers to the thematic group assigned based on keyword co-occurrence. Betweenness, Closeness, and PageRank represent network centrality measures, reflecting a keyword's role as a bridge, its proximity to others, and its overall influence, respectively.

130 Sunghwa Lee · Myoungho Ha · Byong-Rae Ryu · Se-Eun Jhang

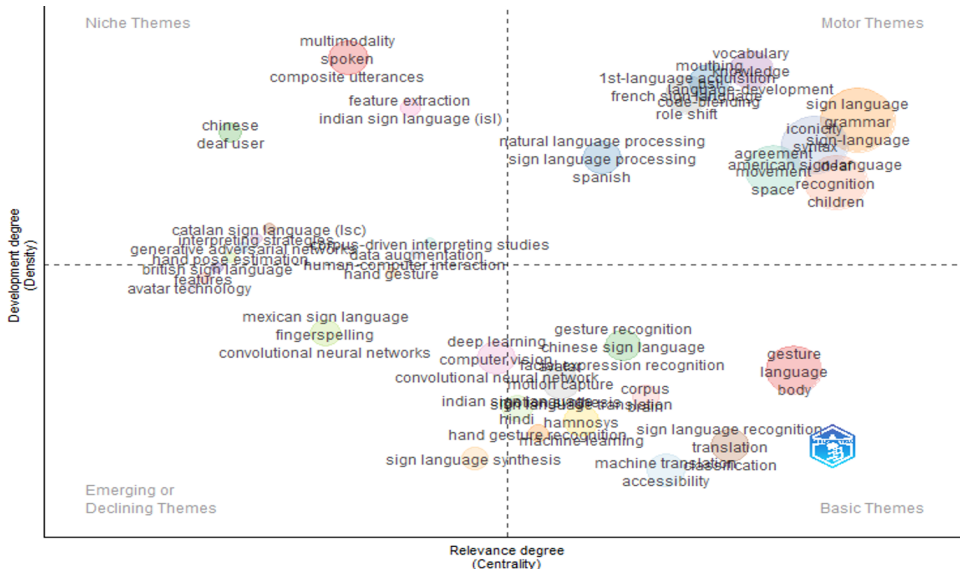


Figure 11. Thematic map of period 3 (2016–2024) for sign language grammar

Figure 11 shows that the thematic space in Period 3 is more cohesive, compared to the fragmented structure of Period 2. Technologically oriented keywords appear prominently in the motor quadrant, suggesting that new research directions are not only emerging but also becoming integrated into the field's core. Topics such as *sign language recognition*, *avatar*, and *machine translation* reflect the increasing role of deep learning and multimodal modeling in grammatical analysis, especially in efforts to develop real-time translation tools and accessible communication technologies for Deaf communities.

At the same time, core concepts such as *sign language* and *language* maintain their central positions, indicating continuity with earlier periods. Their consistent presence, now embedded within a richer technological context, signals a maturation of the field—where foundational themes serve as stable anchors while new research expands outward into interdisciplinary domains.

4.2.2.4 Overall thematic landscape

The overall thematic map of sign language grammar research (Figure 12 below) encapsulates the field's accumulated evolution across all three periods. The most prominent motor cluster is centered on keywords such as *sign language*, *language*, *gesture*, *acquisition*, and *deaf*, indicating the enduring importance of core linguistic constructs alongside increased attention to embodiment and Deaf identity. Notably, the upper-right quadrant—representing motor themes—shows the convergence of foundational grammatical concerns (*language*, *acquisition*) with socially grounded topics like *gesture* and *deaf children*. This suggests that the field has matured to integrate structural analysis with broader communicative and community-based concerns.

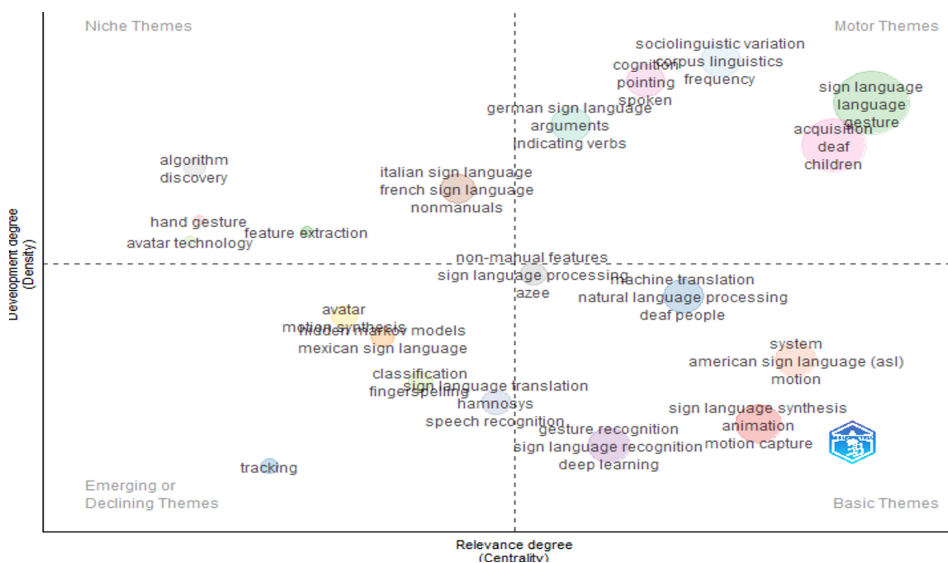


Figure 12. Thematic map of sign language grammar across all periods

In Figure 12, technology-oriented themes such as *machine translation*, *natural language processing*, *sign language recognition*, *deep learning*, and *avatar* appear mostly in the lower-right (basic themes) and lower-left (emerging/declining themes) quadrants. Their presence across these zones suggests they are recognized as relevant but still undergoing structural development and integration. These keywords are crucial for the field's engagement with accessibility and real-time communication tools, but their

dispersion reflects ongoing experimentation and conceptual diversification.

The upper-left niche quadrant contains terms such as *algorithm*, *discovery*, *avatar technology*, and *hand gesture*, indicating specialized subfields with limited centrality. This quadrant reflects more focused, often technologically driven innovations that are not yet widely connected to the field's mainstream discourse.

Altogether, the map illustrates a dual trajectory: while the field continues to reinforce its linguistic foundations, it is also actively branching into new interdisciplinary domains—particularly those intersecting with AI, HCI (human-computer interaction), and assistive technologies. This bifocal thematic structure underscores the maturity and adaptive breadth of sign language grammar research today.

4.3 Persistent keywords across periods

To deepen our understanding of thematic continuity and innovation, this section explores persistent keywords that appeared across all three periods of grammar research in spoken and sign languages. This diachronic keyword analysis traces the conceptual stability of core themes and identifies foundational research directions that have shaped each field over time. Unlike Section 4.2, which examined period-specific keywords, this section adopts a cross-period perspective.

4.3.1 Tracing persistent keywords

Persistent keywords refer to terms that consistently appeared across all three periods. These terms represent enduring research interests and conceptual anchors within the grammar research landscape. Table 4 presents the persistent keywords identified in spoken and sign language grammar studies, including both shared and modality-specific items.

Table 4. Persistent keywords across periods in spoken and sign language grammar research

Modality	Persistent keywords
Spoken languages	acquisition, activation, adults, age, agreement, aphasia, assessment, attention, binding, brain potentials, children, classroom, complexity, comprehension, constraints, constructions, context, conversation, coordination, deficits, discourse, english, eye-movements, french, frequency, grammar, impairment, individual-differences, infants, inflection, information, input, instruction, intonation, japanese, knowledge, language, language development, language impairment, language production, length, lexical access, lexicon, marking, memory, model, morphology, movement, nouns, organization, perception, performance, phonology, pronouns, prosody, recognition, relative clauses, representation, resolution, retrieval, rules, semantics, sentence comprehension, sentence production, sentences, skills, specific language impairment, speech, speech production, stress, students, syntactic complexity, syntax, tense, time, verb, verbs, word, word recognition, words, working memory, working-memory, young-children
Sign languages	second language acquisition, acquisition, agreement, american sign language, children, communication, critical period, deaf-children, english, facial expression, hearing, language, morphology, movement, perception, prosody, recognition, sign language, speech, syntax
Common keywords	acquisition, agreement, children, english, language, morphology, movement, perception, prosody, recognition, speech, syntax

The persistent keywords reveal that grammar research in both modalities maintains strong continuity in terms of core theoretical concerns. These commonly shared keywords can be grouped into several thematic categories:

- (1) Common shared keywords categories.
 - a. Acquisition and development: *acquisition, children*
 - b. Grammatical structure: *agreement, morphology, movement, syntax*
 - c. Processing and production: *perception, prosody, recognition, speech*
 - d. Language type and target: *language, English*

This distribution reflects a comprehensive interest in the entire pipeline of language—from learning to processing and output—across both spoken and sign language research. The presence of keywords spanning multiple levels of linguistic structure (*morphology, syntax, prosody*), cognitive mechanisms (*perception, recognition*), and development (*acquisition, children*) highlights a robust and integrated research orientation.

4.3.2 Language-specific persistent keywords

While a shared core of persistent keywords exists, each modality also demonstrates its own domain-specific emphases, as illustrated below.

(2) Spoken Language keywords categories

- a. Cognitive and processing mechanisms: *attention, working memory, eye-movements, brain potentials, sentence comprehension*
- b. Grammatical and syntactic constructs: *relative clauses, constructions, tense, verb, syntactic complexity*
- c. Language diversity and population: *French, Japanese, infants, young-children*

(3) Sign language keyword categories

- a. Modality-specific and sociolinguistic features: *facial expression, hearing, sign language, American Sign Language, deaf-children*
- b. Development and access: *communication, critical period, second language acquisition*

This contrast indicates that while both fields pursue the common question of how language is learned, processed, and produced, spoken language grammar has evolved with more diversified subfields—anchored in psycholinguistics and formal syntax—while sign language grammar remains strongly oriented around visual modality, language accessibility, and bilingual acquisition. These results provide a conceptual foundation for further reflection. In the following section, we discuss the broader implications of these findings, focusing on thematic convergence, modality-specific orientations, and directions for future linguistic research.

5. Discussion

Building on the results presented in Section 4, this section interprets key findings through a broader conceptual lens. It explores how grammar research in spoken and sign languages has evolved across time, identifies shared foundations and

domain-specific emphases, and considers how these trends inform future linguistic inquiry. In doing so, it draws on both thematic evolution and persistent keyword analyses to assess what has remained stable, what has shifted, and why these patterns matter for the advancement of linguistic theory and methodology.

5.1 Temporal trends and shared evolution

Both spoken and sign language grammar research exhibit a three-phase developmental pattern—Formative Stage, Growth Stage, and Expansion Stage—indicating a broadly shared trajectory of field maturation. In both modalities, early research centered on foundational theoretical constructs such as syntax and acquisition, gradually expanding into studies of processing (e.g., *memory*, *modality*) and more recently into technology-driven themes such as *machine translation* and *recognition*.

This transition reflects a conceptual shift from theory-building toward applied and interdisciplinary domains, particularly in the latest period. The convergence of research interests in technology and cognitive processing across modalities illustrates the increasing relevance of multimodal and AI-based approaches in grammatical studies. This convergence reflects a broader shift from theoretical to applied domains across modalities; however, it also underscores areas where the evolution has been uneven. In particular, the comparatively limited development of syntactic theory in sign language grammar—due in part to visual-spatial constraints—emerges as a gap that warrants deeper exploration, as elaborated in Section 5.3.

5.2 Shared core themes in spoken and sign language grammar

Despite differences in modality, both fields show consistent engagement with a core set of persistent keywords across time: *acquisition*, *agreement*, *children*, *morphology*, *perception*, *prosody*, *recognition*, *speech*, and *syntax*. These recurring terms represent core domains of grammatical inquiry and highlight conceptual continuity over time.

Importantly, the recurrence and co-occurrence of these keywords suggest that grammatical research across modalities is underpinned by an integrated view of language development. This perspective envisions grammar as unfolding through a pipeline of interconnected processes: beginning with acquisition, progressing through

perceptual encoding and recognition, and culminating in language production (spoken or signed). This developmental model offers a useful framework for constructing cross-modality comparisons and may serve as a conceptual bridge between spoken and sign language studies.

These shared themes underscore the possibility of developing integrated processing models that reflect how grammar is acquired, processed, and expressed regardless of modality. By aligning empirical trends across both language types, such models can support the creation of unified theoretical frameworks and inform multimodal research in linguistics, education, and language technology.

5.3 Modality-specific orientations

While the shared core reflects foundational continuity, clear differences emerge in each modality's emphasis. In particular, sign language grammar research emphasizes modality-specific features such as *facial expressions*, *deaf education*, and *accessibility*, focusing heavily on the visual-manual channel and its implications for language learning and processing. Theoretical work in sign language grammar is currently concentrated in morphology and phonetics, with syntax receiving comparatively limited attention. As illustrated in Figure 13, this imbalance likely stems from the visual-spatial constraints of syntactic structure in sign language (Lee & Jhang 2025), which pose challenges to formal modeling.

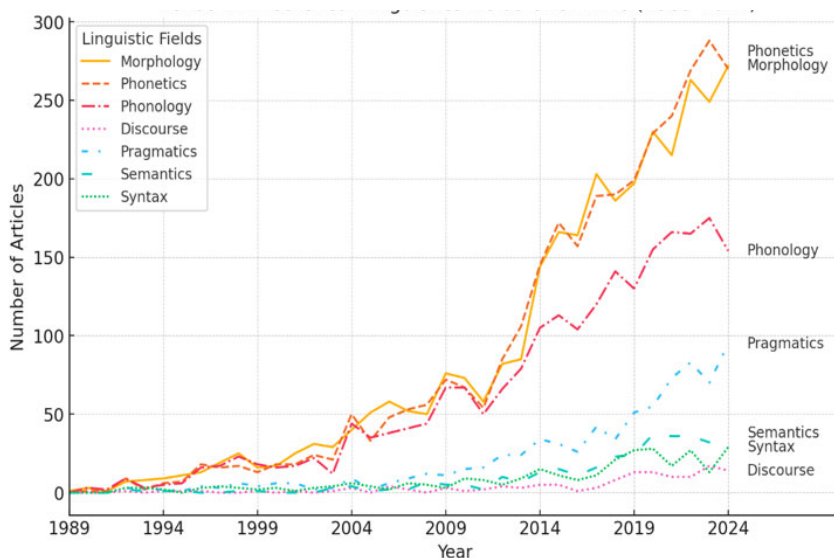


Figure 13. Trends in theoretical subfields of sign language linguistics (1989–2024), highlighting limited growth in syntax relative to morphology and phonetics⁶

In contrast, spoken language grammar research exhibits a more diversified and granular keyword profile. Persistent terms span a wide range of psycholinguistic constructs such as *attention*, *working memory*, *retrieval*, and *sentence comprehension*, pointing to sustained engagement with experimental and cognitive approaches. Additionally, numerous syntactic and morphological constructs—including *relative clauses*, *constructions*, *inflection*, *tense*, *sentence production*, and *syntactic complexity*—indicate a long-standing tradition of detailed structural analysis. The inclusion of keywords like *aphasia*, *impairment*, *young-children*, and *French/Japanese* further demonstrates the field's attention to clinical, developmental, and cross-linguistic variation. Collectively, these patterns reflect spoken language grammar's strong roots in formal theory, neurolinguistics, and language acquisition research, supported by extensive empirical methodologies such as eye-tracking, ERP, and reaction-time paradigms.

These contrasts in focus and methodological scope underscore how modality influences linguistic inquiry. This divergence highlights how modality shapes the

⁶ Adapted from Lee and Jhang (2025).

direction and depth of grammatical inquiry. In sign language research, visual-spatial constraints have steered theoretical focus toward morphology and phonetics, while limiting syntactic exploration. In contrast, the auditory-verbal modality of spoken language has enabled detailed modeling of syntactic and psycholinguistic processes within well-established theoretical paradigms. These differences not only reflect distinct challenges but also reveal untapped opportunities for developing modality-sensitive linguistic frameworks.

5.4 On the issue of normalization in unequal datasets

One methodological concern in this study lies in the marked imbalance between the spoken language dataset (22,947 articles) and the sign language dataset (712 articles). To mitigate this 32-fold disparity, normalization techniques such as subsampling, frequency adjustment, or bootstrap resampling could have been employed. However, normalization was deliberately not applied here. The rationale was that the purpose of this research was not to artificially equalize the two datasets, but rather to demonstrate how differences in research productivity and thematic diversity are reflected in the bibliometric landscape.

Had normalization been carried out, the findings would likely have differed in several respects. First, in methodological terms, subsampling or frequency scaling would have reduced the size of the spoken language corpus, allowing for more direct quantitative comparisons with sign language. Second, in terms of results, the number and size of clusters in spoken language research would have appeared smaller, more closely resembling the compact structures observed in sign language research. Likewise, the number of persistent keywords identified in spoken language would have decreased, narrowing the apparent gap between the two modalities. Third, in terms of analysis, normalization would have emphasized thematic similarities between spoken and sign language grammar while downplaying the substantive differences in scope and diversity that in fact characterize the two fields.

By refraining from normalization, this study highlights the empirical reality that spoken language grammar research, supported by a longer tradition, larger data pools, and stronger institutional infrastructure, has diversified into a broader thematic landscape. In contrast, sign language grammar research remains centered on a smaller

set of modality-specific concerns. While this decision limits the extent to which the two datasets can be compared under strictly equal statistical conditions, it provides a more accurate reflection of the structural disparities between the fields. For the sake of transparency, we acknowledge this trade-off as a limitation. Future work may complement the present analysis by applying normalization or resampling techniques to assess the robustness of the patterns observed here.

5.5 Limitations of the study

In addition to dataset imbalance, several broader limitations should be acknowledged. First, while co-word analysis is useful for mapping thematic structures, its reliance on keyword co-occurrence can overemphasize high-frequency terms and fail to capture finer conceptual nuances. Second, although reliance on the Web of Science Core Collection ensured consistency and quality, it necessarily excluded relevant works indexed in other databases such as Scopus or KCI, thereby limiting comprehensiveness. Third, the analysis focused solely on English-language publications, which may have left region-specific or non-English scholarship—particularly in sign language research—underrepresented. Finally, the substantial disparity in dataset size between spoken language (22,947 articles) and sign language (712 articles) directly influenced cluster stability and the criteria applied for periodization.

Nevertheless, these limitations do not diminish the validity of the main findings. Rather, they point to avenues for future research, such as incorporating multiple databases, multilingual corpora, and complementary analytic methods to extend the present study.

5.6 Implications and future directions

While this study offers a detailed comparative analysis of grammar research in spoken and sign languages through thematic evolution, it also acknowledges the inherent limitations of bibliometric methods. Bibliometric analyses are constrained by the availability and structure of publication metadata, often lacking access to the full semantic depth of individual studies. Moreover, the interpretive accuracy of keyword-based mapping depends on the consistency and clarity of author-defined

terms.

Despite these constraints, the study demonstrates that a well-structured bibliometric approach—particularly when coupled with careful thematic segmentation and persistent keyword analysis—can yield meaningful insights into disciplinary trends and research gaps. By doing so, it not only traces the development of linguistic inquiry across modalities but also identifies critical areas for theoretical and methodological advancement, especially within the sign language domain. The following subsections outline four key directions informed by these findings.

5.6.1 Toward multimodal grammar models

The persistence of shared keywords such as *acquisition*, *perception*, *recognition*, and *speech* across both modalities points to the potential for developing integrated models of language processing. These models could map out a full linguistic pipeline—from learning and comprehension to production—while incorporating modality-specific features. For instance, sign language grammar requires the integration of visual-spatial structures such as *spatial indexing*, *simultaneity*, and *non-manual markers*, which pose challenges to linear, speech-based models of grammar. In contrast, spoken language relies more heavily on sequential auditory cues and prosodic structuring. A unified framework would need to reconcile these modality-specific features through abstract representations that are modality-independent, such as conceptual dependency structures or multimodal syntax trees. Such a model could serve as a foundation for developing cross-linguistic processing systems, enhancing sign language translation technologies, and informing inclusive educational interventions that accommodate diverse language modalities.

5.6.2 Balancing theoretical development in sign language

As discussed in Section 5.3, syntactic theory in sign language remains underdeveloped compared to other subfields like morphology and phonology. This gap stems in part from the unique challenges posed by visual-spatial structures and non-linear syntax in sign languages. Addressing this imbalance requires adapting or extending formal grammatical frameworks such as generative grammar or constraint-based models to

account for spatial indexing, non-manual signals, and simultaneity. Advancing syntactic modeling would enhance the theoretical completeness of sign language linguistics and improve cross-modal comparisons.

5.6.3 Expanding psycholinguistic research in sign language

Spoken language grammar research has made extensive use of experimental psycholinguistic methodologies, as reflected in persistent keywords like *attention*, *working memory*, *retrieval*, and *sentence comprehension*. In contrast, sign language studies have yet to fully incorporate such approaches. Greater adoption of techniques such as eye-tracking, ERP, and behavioral experiments would help uncover the cognitive processes underlying sign language use, offering new insights into language processing across modalities and contributing to neurolinguistic theory more broadly.

5.6.4 Promoting cross-modality comparative research

The divergence in persistent and emerging keywords across modalities underscores the value of systematic comparative studies. Future research should aim to determine whether observed differences stem from modality, linguistic structure, or disciplinary tradition. Interdisciplinary collaboration between researchers in spoken and sign languages—spanning theoretical linguistics, psycholinguistics, and computational modeling—can enrich both fields. Importantly, theoretical insights from sign language studies may inform spoken language research in return, fostering a more integrated and inclusive understanding of human grammar.

6. Conclusion

This study explored the evolution of grammatical research in spoken and sign languages using bibliometric methods, with a focus on thematic trajectories and persistent conceptual anchors. By analyzing over three decades of literature, we identified both shared and modality-specific patterns that define the current state and future directions of linguistic inquiry.

First, both spoken and sign language grammar studies followed a parallel

three-stage development—formative, growth, and expansion. Early research emphasized theoretical constructs such as syntax and acquisition, which later gave way to experimental paradigms and, more recently, to technology-driven topics including machine translation and AI-based recognition. This shared arc reflects a convergent progression toward applied and interdisciplinary engagement.

Second, the analysis of persistent keywords revealed a stable conceptual foundation across modalities. Recurring terms such as *acquisition*, *agreement*, *syntax*, *speech*, and *recognition* point to enduring concerns with how language is acquired, processed, and produced. These anchors provide a strong empirical basis for developing comparative and multimodal models of grammar.

Third, each modality demonstrated distinct thematic orientations. Sign language grammar research concentrated on modality-specific themes such as *facial expression*, *deaf education*, and *accessibility*, while formal syntactic theory remains relatively underexplored. In contrast, spoken language grammar has developed a broader and more detailed engagement with psycholinguistics, neurolinguistics, and typological variation, supported by diverse empirical methodologies.

Several gaps in sign language grammar research were identified, including (i) the need for comprehensive models linking processing and production, (ii) the integration of experimental techniques such as ERP and eye-tracking, and (iii) the development of objective metrics for production fluency and accuracy. These areas offer immediate opportunities for interdisciplinary collaboration and theoretical advancement.

By moving beyond descriptive metrics and incorporating interpretive thematic analysis, this study sheds light on the developmental trajectories and conceptual anchors of both spoken and sign language grammar. In doing so, it reveals not only the shared foundations of grammatical inquiry but also the distinctive pathways shaped by modality. As sign language research continues to gain momentum, addressing its underrepresented areas—such as syntax, psycholinguistics, and production modeling—will be key to establishing a more complete and autonomous field of grammatical theory. This study offers an empirical foundation for that effort, pointing the way toward more systematic, multimodal, and theoretically grounded approaches to sign language grammar.

Building on the present findings, future research could further explore the thematic architecture of grammar studies by applying topic modeling and network analysis to the same datasets. Such approaches would offer a more fine-grained view of

conceptual linkages and topic clusters, complementing the bibliometric trends identified here and advancing a deeper understanding of how grammatical research evolves across modalities.

References

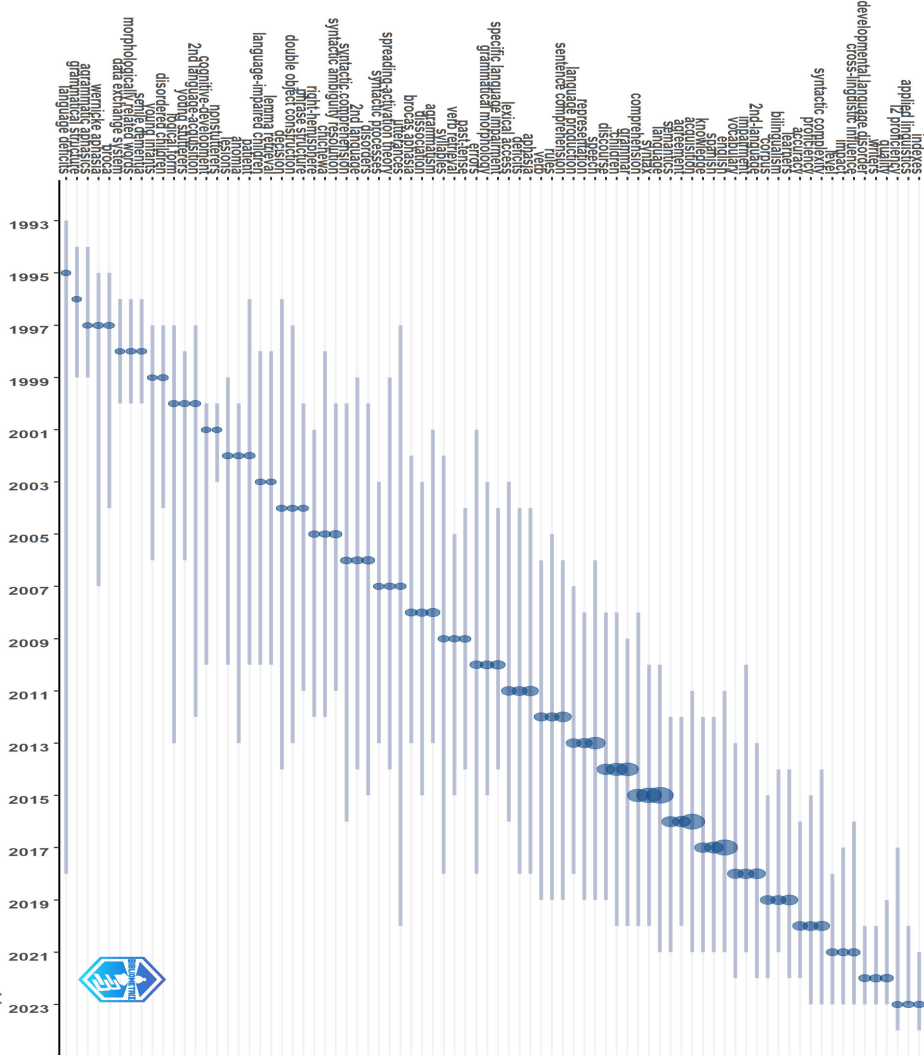
- Aria, Massimo and Corrado Cuccurullo. 2017. Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11(4): 959–975.
- Arik, Egin. 2014. Sign language research in Web of Science. *Journal of Scientometric Research* 3(3): 143–149.
- Bresnan, Joan. 2001. *Lexical-functional syntax*. Oxford: Blackwell.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Crosthwaite, Peter, Sumiati Ningrum, and Martin Schweinberger. 2023. Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities. *International Journal of Corpus Linguistics* 28(3): 344–377.
- Donthu, Naveen, Satish Kumar, Debidutta Mukherjee, Nitesh Pandey, and Weng Marc Lim. 2021. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* 133(5): 285–296.
- Hajar, Assim and Muhammed Karakus. 2025. Five decades of language learning strategy research: A bibliometric review and research agenda. *The Language Learning Journal* 53(2): 220–249.
- Honnibal, Matthew, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. *spaCy: Industrial-strength natural language processing in Python*. Zenodo.
- Lee, Sunghwa and Se-Eun Jhang. 2025. Mapping the linguistic landscape of sign language: Research trends and thematic insights. Paper presented at *The 2025 Winter Conference of the Korean Society of Linguistic Science*. Busan, Pusan National University, February 19.
- Lillo-Martin, Diane and Richard P. Meier. 2011. On the linguistic status of ‘agreement’ in sign languages. *Theoretical Linguistics* 37(3-4): 95–141.
- Liu, Cuilin. 2025. *Integrating bibliometric techniques with corpus-driven bigrams: A research trend analysis of cruise tourism*. Phd Dissertation. National Korea Maritime & Ocean University.
- Liu, Cuilin and Se-Eun Jhang. 2024. A bibliometric analysis of language acquisition research: A comparative study of spoken and sign languages. Paper presented at *The 2nd Workshop on Interdisciplinary Research on Acquisition of the Korean Sign Language*. Chungnam National University, December 13.
- Liu, Cuilin, Ziyun Dai, and Se-Eun Jhang. 2024. Bibliometric analysis of research on gender equality in the maritime industry. *The New Korean Association of English Language and Literature* 66(3): 171–199.

- Narenthiran, Rajagopal, Govindaraj Saravanan, and Sathi Sa Aswathy. 2016. Sign language research 1991–2015: A bibliometric study with Web of Science. *National Seminar on Librametrics to Altmetrics (NSLA 2016)*, Department of Library and Information Science, University of Kerala, Feb 2016, Thiruvananthapuram, India.
- Parlina, Asmi, Kiki Ramli, and Herman Murfi. 2020. Theme mapping and bibliometrics analysis of one decade of big data research in the Scopus database. *Information* 11(2): 69.
- Perianes-Rodríguez, Antonio, Ludo Waltman, and Nees Jan Van Eck. 2016. Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics* 10(4): 1178–1195.
- Sandler, Wendy and Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.
- Van Eck, Nees Jan and Ludo Waltman. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84(2): 523–538.
- Waltman, Ludo. 2016. A review of the literature on citation impact indicators. *Journal of Informetrics* 10(2): 365–391.
- Zupic, Ivan and Tomaž Čater. 2015. Bibliometric methods in management and organization. *Organizational Research Methods* 18(3): 429–472.

Appendices

Appendix A

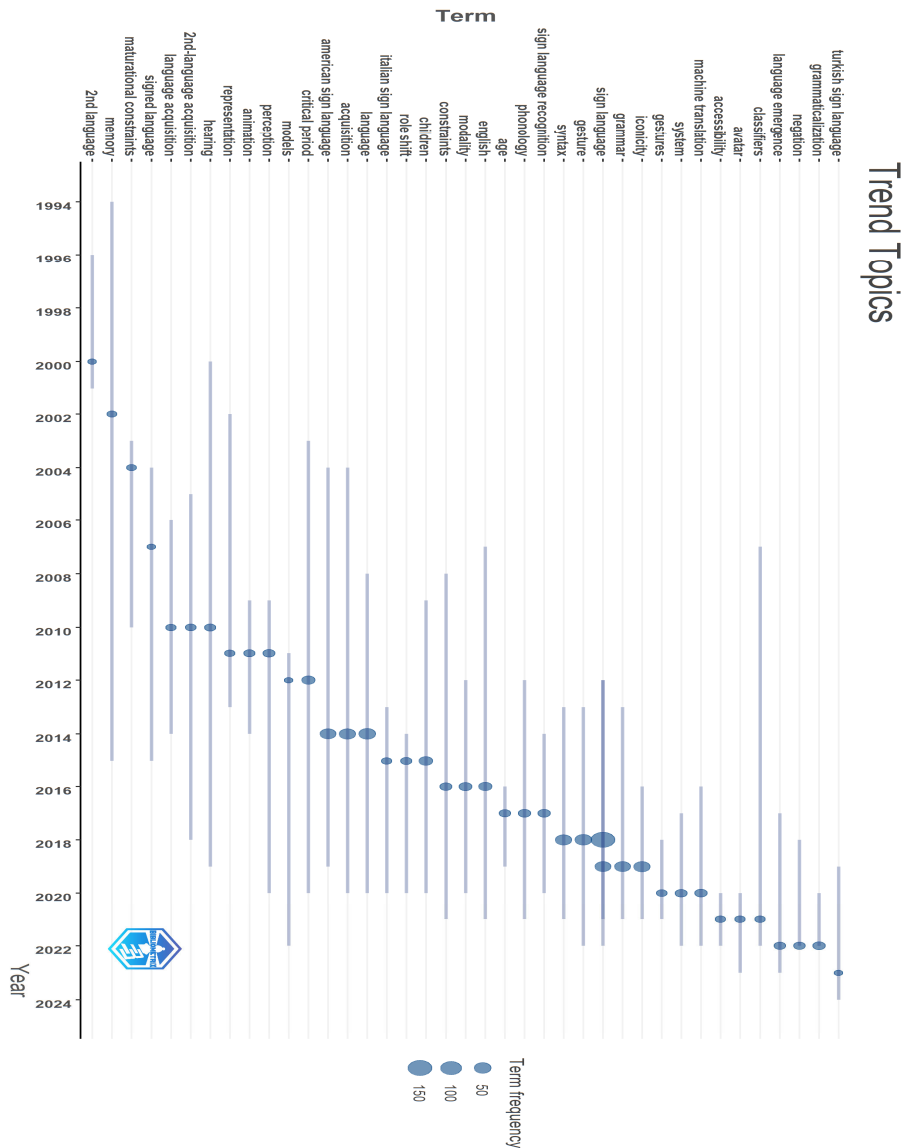
Figure 2: Trend topics for spoken language grammar research (1990–2024)⁷



⁷ The size of each dot indicates the frequency of the keyword; horizontal lines indicate the active period of each keyword.

Appendix B

Figure 4. Trend topics for sign language grammar research (1989–2024)



Sunghwa Lee

Research Professor
Department of English Language and Literature
National Korea Maritime & Ocean University
727 Taejong-ro, Yeongdo-gu,
Busan 49112, Korea
Email: eabbysunghwa@gmail.com

Myoungho Ha

Assistant Professor
Silla University
140 Baegyang-daero(Blvd), 700beon-gil(Rd),
Sasang-gu, Busan 46958, Korea
Email: hadash21@silla.ac.kr

Byong-Rae Ryu

Professor
Department of Linguistics
Chungnam National University
99 Daehak-ro, Yuseong-gu,
Daejeon 34134, Korea
Email: ryu@cnu.ac.kr

Se-Eun Jhang

Professor
Department of English Language and Literature
National Korea Maritime & Ocean University
727 Taejong-ro, Yeongdo-gu,
Busan 49112, Korea
Email: jhang@kmou.ac.kr

Received: 2025. 08. 05.

Revised: 2025. 09. 12.

Accepted: 2025. 09. 17.