SCIENTIFIC Publishing Limited

Neurolinguistic Communication Intervention

https://ojs.ukscip.com/index.php/nci

Review

Retiring Linguistics for a Unified Language Science

Irina Albu [1,2,*] and Torben Edward Voß [1]

[1] Department of Psychology, Education, and Child Studies, Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, 3062 PA Rotterdam, The Netherlands

[2] Faculty of Arts, Humanities and Social Sciences, School of Psychology, Trinity College Dublin, D02 PN40 Dublin, Ireland

[*] Correspondence: 686826ia@student.eur.nl

Received: 3 December 2025; **Revised:** 7 January 2026; **Accepted:** 22 February 2026; **Published:** 11 March 2026

Abstract: Language research has never been richer; spanning formal theory, documentation, neuroscience, psychology, education and AI. Yet it remains partitioned by disciplinary silos, methodological habits, and Western, Educated, Industrialized, Rich, and Democratic (WEIRD) sampling biases. We argue that sustained progress on core problems, i.e., how language is learned, processed, varies, breaks down and can be engineered, requires "retiring linguistics" as an isolated discipline and consolidating expertise within an integrated Language Science. This shift does not compromise rigor; rather, it situates formal modelling alongside quantitative evidence, field-based research, clinical and technological applications. Recent advances that bridge traditional boundaries such as neurosemantic mapping, speech neuroprosthetics, and computational approaches to sociolinguistic variation illustrate the potential of such integration. At the same time structural barriers including departmental incentives, fragmented training pathways, and inconsistent terminology continue to limit coordination across fields. This article combines a critical review of contemporary language research with a concrete proposal for institutional and epistemic reforms: transdisciplinary institutes and appointments, evaluation criteria that reward collaboration, curricula that braid theory, computation and field methods, funding and venues for cross-field work, ethical frameworks centred on partnership, and benefit-sharing with language communities. Unifying around problems rather than on departments can deliver more universally applicable research and greater societal benefit, from equitable language technologies and education, to improved clinical outcomes, by aligning explanations across levels from neurons to social networks.

Keywords: Language Science; Interdisciplinary Linguistics; Psycholinguistics; Methodological Integration; Linguistic Diversity

1. Introduction

Linguistic research has yielded transformative insights into the structure and diversity of human language, from formal grammar and phonological models to typological surveys and documentation of thousands of languages. Classical linguistics famously identified universal patterns (e.g., syntactic hierarchies, semantic roles) and described innate cognitive capacities for language often described as a "language instinct" [1,2]. At the same time, evidence of enormous linguistic diversity [3] and usage-driven phenomena has prompted a paradigm shift toward usage-based, statistical, and ecological models of language.

Usage-based approaches gained prominence, because they provide principled explanations for empirical patterns that are difficult to reconcile with strictly formal models of grammar. Frequency effects show that repeated ex-

posure leads to entrenchment, increased productivity, and structural change [4]. Linguistic categories often display gradient membership consistent with probabilistic learning mechanisms that yield graded generalisations rather than discrete symbolic classes [5]. Within this framework, Construction Grammar conceptualises linguistic knowledge as a network of learned form-meaning pairings whose productivity and stability are shaped by competition, frequency, and communicative function [6]. Developmental research supports this view by demonstrating that children initially acquire item-based schemas embedded in specific usage contexts and only gradually abstract broader patterns through domain-general learning processes [7]. These usage-based representations align with psycholinguistic and computational findings that show that processing constraints such as predictability, memory limitations, and ambiguity avoidance systematically influence which linguistic structures become stable and frequent across languages [8].

These achievements form a rich foundation, yet the traditional discipline of "linguistics" is struggling to capture the full complexity of language as it is studied today across numerous fields. What we provocatively call "retiring linguistics" would mean dissolving linguistics as an isolated academic silo while preserving its knowledge. The proposal is not to abandon linguistic expertise, but to develop it into a broader, pluralistic Language Science that integrates formal, cognitive, computational, social, clinical, and technological approaches. An integrated approach would work towards tighter cooperation between theory-driven or formal approaches to language and their experimental and empirical counterparts. This would establish evidence that is easier to interpret across the science as a whole, as well as in broader conjunction of linguistics as a field of study with other disciplines. As such, we argue not for a straightforward abolishment of linguistics as a discipline, but for institutional reforms that handle the current structural barriers resulting from current linguistic practice.

Language Science is conceived as a holistic meta-discipline. For example, the University of Maryland Language Science Centre explicitly "advances an interdisciplinary science of language" that spans departments and connects fundamental science to learning, technology, and clinical applications. At Penn, the ILST initiative similarly aims "to integrate the computational, structural, cognitive, and neuroscientific study of speech, language and communication" across the university [9]. Such programs embody the idea that language science subsumes linguistics, neuroscience, psychology, AI, anthropology, education, and more, united by common goals. The ultimate vision is that language research should be organised not by discipline, (comp-sci vs. psych vs. anthropology) but by problems and phenomena (speech perception, semantics, discourse, language learning, language disorders, communication technologies).

2. Why Linguistics Cannot Stand Alone: Language as a Multi-Level Phenomenon

A growing body of cognitive science research supports the view that language is implemented in a distributed brain architecture that cannot be cleanly isolated from domain-general prediction, memory, and control systems, making a purely "linguistics-internal" account incomplete. Work on naturalistic comprehension emphasises that language understanding is built on general neurocomputational principles (e.g., predictive processing and multi-scale integration of context) rather than language-specific mechanisms alone [10]. Consistent with this, neural and computational studies show that both brains and deep language models rely on continuous prediction and context-dependent representations during narrative processing, suggesting that the relevant explanatory targets include general principles of sequence modelling and prediction, not just symbolic grammatical constraints [11]. At the same time, careful functional-localisation research argues that there is a selective frontotemporal "language network" that is robust across individuals, but crucially this network sits within a broader landscape of specialised and domain-general systems that jointly enable language behaviour [12]. Large-scale functional magnetic resonance imaging (fMRI) evidence further refines this picture by showing that classic executive-control circuitry (the multiple-demand network) is not necessary for core comprehension in passive conditions yet becomes engaged when comprehension is embedded in task demands. This finding highlights that language use in real life routinely recruits interacting systems beyond grammar proper [13]. Taken together, the emerging theoretical implication is not "no specialisation," but no single-discipline sufficiency; explaining language requires jointly modelling specialised linguistic representations and the domain-general systems that enable prediction, integration, and goal-directed use in context.

Another reason "retiring linguistics" becomes methodologically and theoretically compelling is a persistent levels-of-analysis mismatch. Traditional linguistic description excels at characterising structural regularities yet of-

ten does not specify how those structures are implemented in neural dynamics, learned under memory constraints, or realised in interaction. Naturalistic neuroscience has made this mismatch explicit by showing that language comprehension unfolds across multiple timescales and interacts with non-linguistic signals, making it hard to map neat theoretical units (e.g., "syntax" or "semantics") onto separable neural subprocesses without integrative models [14,15]. Theoretical proposals in cognitive science likewise argue that fleeting memory creates a "now-or-never" constraint, forcing the system to rapidly compress and recode input, implying that linguistic structure must be compatible with real-time processing and learning instead of mere static well-formedness [16]. In parallel, frameworks that aim to "ground the neurobiology of language in first principles" emphasise computation, context, and prediction as organising axes, again pushing against explanations that stay confined to a single descriptive level [17]. The upshot is that cumulative progress increasingly depends on bridging structure ↔ processing ↔ learning ↔ neural implementation ↔ interaction, which is precisely the kind of cross-level integration that a pluralistic Language Science is designed to institutionalise rather than treat as "adjacent" work.

Modern research views language as a dynamic, adaptive system shaped by feedback loops between cognition, interaction, and cultural transmission. Therefore, the most holistic explanatory theories must be integrative, combining insights across disciplines. Cultural-evolution models (including iterated learning and population-structure approaches) show how key properties of language can emerge and stabilise because they improve learnability and transmission [18]. Recent experimental-cultural work demonstrates that core statistical features of language (e.g., part-like segmentation and distributional regularities) can arise across generations and become increasingly learnable over time. This provides mechanistic evidence that language structure is jointly a product and a driver of cultural dynamics [19]. Complementing this, information-theoretic approaches frame cross-linguistic regularities as outcomes of pressures for efficient communication. They explicitly position information theory as a bridge between formal descriptions and functional or processing explanations [20]. Importantly, the "ecological" turn remains an active empirical debate rather than a settled conclusion. Recent work testing whether socio-demographic variables (e.g., speaker population, L2 learning ecology) reliably predict structural simplification produces mixed results [21]. This demonstrates the need for integrative causal models rather than single-factor stories. Overall, if language is simultaneously (i) neurally implemented in real-time prediction and multiscale integration, and (ii) culturally evolved under learnability and efficiency pressures, then maintaining "linguistics" as an isolated silo becomes increasingly untenable because the causal explanatory machinery naturally crosses disciplinary boundaries.

3. Current Landscape: Fragmentation in Contemporary Language Research

Despite this conclusion, language inquiry remains fragmented across many fields. As one commentator observes, "the study of language has fragmented into many highly specialised areas of study that tend not to talk to each other." In practice, experts in education, engineering, neuroscience, field linguistics, clinical sciences, policy, and AI receive different training, are housed in different departments, publish in different venues, and often have little sustained contact. As a result, even foundational questions about what language is, how it is processed, and how it develops become siloed rather than cumulative.

This fragmentation manifests both between fields and within linguistics itself. Researchers working on language evolution frequently disregard developmental and processing data; studies of child language acquisition may ignore evolutionary constraints; psycholinguists often focus on adult processing in controlled laboratory contexts without full engagement with social or developmental variation. Computational linguists developing artificial intelligence (AI) and natural language processing (NLP) systems frequently treat linguistic structure as secondary to statistical performance, while sociolinguists and anthropological linguists investigate language in society with minimal reference to formal models or brain mechanisms. Education specialists grapple with literacy and classroom language challenges largely separately from cognitive and psycholinguistic models of reading, and clinical fields such as speech-language pathology may draw selectively on phonetic or cognitive models without engaging the full breadth of linguistic theory. Even within linguistic subdisciplines, different frameworks often share few central assumptions. Some examples are contrasts between phrase-structure and dependency grammars, or between constraint-based and Optimality Theory (OT) approaches and frameworks that do not adopt OT, Universal Grammar and frameworks that contest it. Consequently, linguistics struggles to function as a cumulative science, because incompatible assumptions make it difficult to interpret claims across linguistics as a field.

Empirical evidence for this siloing is striking. Large-scale reviews in psycholinguistics and cognitive science

document severe sampling bias, with studies overwhelmingly focusing on English-speaking, North American university students, while languages and populations from the Global South remain severely underrepresented. This WEIRD bias [22] has been repeatedly identified as distorting generalisations about language and cognition yet persists despite sustained critique over the last decade [23]. Citation analyses further show that cross-disciplinary uptake remains limited even when research questions overlap substantively, and genuine interdisciplinary journals or conferences continue to be the exception rather than the norm.

Similar problems arise in language documentation. While scholars emphasise that comprehensive documentation necessarily requires integration across ecology, anthropology, cognition, and pedagogy, many documentation projects remain narrowly descriptive. As a result, they are often only marginally useful to other subfields such as psycholinguistics or acquisition, limiting their broader scientific impact. In short, language is central to society, cognition, and technology, yet researchers studying it frequently do not share methods, data, or theoretical commitments leading to redundant, incompatible, or non-cumulative efforts.

These structural divisions translate directly into fragmentation in practice where different subdisciplines produce mutually exclusive or contradictory theories that are rarely reconciled. Examples of this can be found in formal syntax, particularly generative syntax, where the concept of Universal Grammar (UG) is contested by linguists from other silos. Critics of the concept argue that the assumption of universality underlying natural languages at a more abstract level is not adequately justified. A common candidate for a universal property among human languages is constituency (and the possible recursion of constituents). Yet, serious cross-linguistic evidence has been brought forward against this view, aspiring to show that recursion may be dispensable in some natural languages, or that embedding traditionally conceived to happen at the level of syntax may actually take place in the morphological domain. The disagreement on such foundational assumptions presents a serious structural barrier to cooperation between silos [3]. But disagreement does not only persist at the foundational level. Linguists may also disagree sharply on the concrete mechanisms that drive syntactic phenomena or the status of syntax in human languages. These disagreements may concern completely separate methodologies. An example of this concerns Phase Theory [24]; a foundational mechanism in the Minimalist Program (MP), a framework of generative syntax. Phase Theory posits that syntactic derivations are built in discrete units ("phases"), imposing strict locality constraints on syntactic operations. This assumption remains central to contemporary syntactic theory and continues to be defended and elaborated in recent work [25,26]. However, psycholinguistic and neurocognitive evidence increasingly challenges the psychological plausibility of strict phase-based locality. Electroencephalography (EEG) and reading-time studies indicate that sentence comprehension relies on predictive processing, global tree-scaffolding, and graded processing costs rather than discrete derivational cutoffs. Studies show that readers maintain access to non-local syntactic information and engage in anticipatory structure building that appears incompatible with strict phase boundaries [27,28]. Related work in sentence processing demonstrates that comprehenders routinely violate locality constraints predicted by narrow syntactic derivations in favor of probabilistic expectations [29]. Despite this accumulating evidence, Phase Theory largely continues to operate independently of processing research, illustrating that a disconnect between formal syntax and psycholinguistics may make it difficult to apply or interpret insights produced by phase-theoretic syntax. However, there are also cases of successful cross-silo cooperation. This concerns for example other aspects of generative syntax, like the view that syntactic processes precede semantic evaluation. Evidence from fMRI studies has argued for compatibility of neurological and psycholinguistic evidence with syntax-first assumptions and aspects of phrase structure building in the brain, constituting a positive example of using psycholinguistic methodology to provide evidential support to phrase-structure grammar [30].

A comparable disconnection exists between formal pragmatics and psycholinguistics. While both traditions aim to explain meaning construction in context, they have increasingly diverged in methods and assumptions. Formal pragmatics typically relies on idealised speaker–listener models and logical inference, whereas psycholinguistics emphasises real-time processing constraints, memory limitations, and prediction under uncertainty. Although some integrative proposals exist, the two fields have largely grown apart [31]. Experimental research demonstrates that pragmatic inference is strongly shaped by cognitive load, expectations, and probabilistic reasoning [32], yet these findings are rarely incorporated into formal pragmatic models.

There are, however, broader disagreements that do not concern psycholinguist methodology but instead reflects possibly irreconcilable theoretical commitments between silos. One of the most longstanding and illustrative cases of such fragmentation persists between phonetics and phonology, particularly in debates over incomplete

neutralisation (IN). Phonetic accounts of final devoicing in German argue that neutralisation is gradient and incomplete, such that devoiced plosives remain measurably distinct from underlyingly voiceless ones [33]. In contrast, phonological accounts maintain that phonological categories remain discrete, interpreting observed phonetic differences as enhancement or analogy rather than genuine phonological contrast. This disagreement continues to generate new work [34] yet persists largely because each subdiscipline employs different criteria for what constitutes a meaningful linguistic distinction. The conflict is sustained less by disagreement over empirical facts than by incompatible methodological and ontological assumptions.

Fragmentation is intensified in the relationship between linguistics and AI/NLP. Contemporary NLP systems achieve impressive performance through scale and statistical learning, often without explicit linguistic representations. While this has driven rapid technological progress, it has also widened the gap between computational approaches and linguistic theory. Empirical analyses show that neural language models can approximate surface-level syntactic patterns while failing to robustly represent hierarchical dependencies central to linguistic theory [35]. Linguistic insights are frequently treated as post-hoc interpretive tools rather than constraints on model design, reinforcing siloed development.

Taken together, these examples demonstrate how fragmentation across and within language-related disciplines produces mutually exclusive accounts with limited opportunity for reconciliation. The divisions extend beyond empirical disagreement to encompass divergent methodologies, explanatory goals, and assumptions about what counts as evidence. As a result, language research advances along parallel tracks with limited cumulative synthesis. Given the central role of language in cognition, education, clinical practice, and technology this siloing represents a structural obstacle to developing a coherent and integrative science of language. We argue that while these distinctions may not be problematic in their respective silos, they may considerably inhibit meaningful interpretation across domains, leaving linguistics with insights that are hard to reconcile with each other. Notably however, there already exist present examples of successful cross-faculty integration, as between psycholinguistics and formal syntax regarding phrase structure. These are positive examples of what an integrated language science with tighter cooperation would encourage and work towards.

4. Structural Barriers and Incentives

These disconnections are perpetuated by structural and institutional factors. Universities are organised into departments that act as intellectual "silos"; each discipline develops its own publication outlets, funding channels, and evaluation criteria. In general, the field has been characterised by a distinction between two perspectives, CL1 (holistic) and CL2 (modular), each with its own foundational assumptions and methodological preferences. This division can limit cross-paradigm dialogue, as researchers tend to work within established frameworks rather than integrating insights from alternative approaches or external evidence [36]. In such an environment, collaboration across fields faces headwinds. As a possible empirical indicator of such fragmentation, we have conducted a brief cross-citation analysis using the OpenAlex library [37]. The results are presented in **Figure 1**. Across 50 articles from each of five flagship journals for relevant linguistic topics, *Linguistic Inquiry*, *Language*, *Natural Language & Linguistic Theory*, *Brain and Language* and *Cognition*, we identified over 10,000 references. After categorizing these journals into academic subfields and examining the cross-references between them, we find asymmetrical citations across these fields, with Formal Syntax appearing to be strongly isolated, citing the other journals only ten times combined.

Other issues concern interdisciplinary cooperation more directly. Reviewers often apply a disciplinary lens to interdisciplinary proposals. Departments primarily hire and reward specialists in traditional areas, so researchers who cross boundaries risk being seen as unfocused or "not real scientists" in any one field. Undergraduate and K-12 curricula contribute too, as language is often taught as a humanities subject obscuring its connections to science and technology.

Recent empirical research on evaluation systems supports this general mechanism. Observational and interview-based work on grant-panel dynamics shows that interdisciplinary proposals can face systematic friction. This occurs because panelists struggle to match them to familiar standards of "quality," often requesting mono-disciplinary framing or penalising category-spanning contributions [38]. Related work on evaluation across fields finds that interdisciplinary work can experience inconsistent outcomes depending on which evaluative cues dominate (topic fit vs. knowledge-base breadth). This reinforces the claim that "disciplinary lenses" function as gatekeeping mech-

anisms with practical consequences [39]. In parallel, research on hiring, promotion, and tenure decision-making shows that evaluators frequently fall back on conventional metrics and category norms when judging "non-standard" scholarship. This makes boundary-spanning language science risky even when institutions verbally endorse interdisciplinarity [40,41].

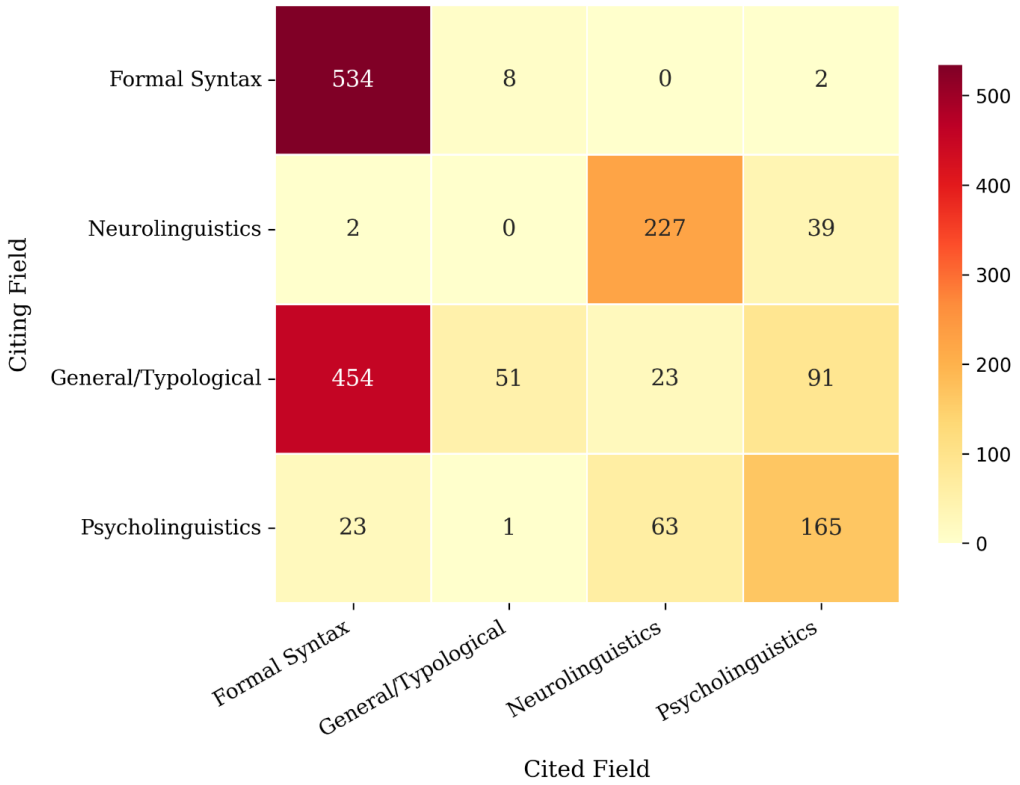

Figure 1. Cross-Citations across linguistic Journals.

There is a considerable lack of convention on what the object and nature of linguistics as a science entails, particularly with regards to formal syntax. While some linguists, such as Noam Chomsky, treat linguistics as part of cognitive psychology and argue that syntactic principles reflect covert mental processes, linguistic research in practice often diverges from this view. In many cases, it remains focused on the study of abstract linguistic entities. As a result, it does not always constitute an empirical investigation of speakers' cognitive faculties and competence [42]. On the other hand, psycholinguistic inquiry often disregards the abstract properties of syntactic and linguistic structures in general. This "object-of-study mismatch" is intensified by the way subfields operationalise evidence. Formal syntax prioritises internal explanatory elegance (e.g., derivational economy and constraint interaction); psycholinguistics focuses on measurable behavioural signatures under task constraints; and natural language processing emphasises benchmark performance under data constraints. When the institutional reward structure treats these as separate career tracks, the mismatch becomes self-reinforcing rather than self-correcting.

Methodologically, fragmentation persists. Some scholars emphasise quantitative, corpus-based methods (as in AI/NLP and corpus linguistics), while others favour qualitative or ethnographic approaches (common in sociolinguistics and education). Many linguistic subfields have evolved their own techniques. For example, in formal syntax, introspection remains a dominant method of inquiry about grammaticality, despite considerable disadvantages [43] and a possibility for linguists to disagree with regards to their idiolects on the grammaticality of a syntactic structure, limiting the generalisability of results produced. As demonstrated by Linzen and Oseki, judgment reliability varies substantially across languages and phenomena, and relying on small-N expert introspection can produce unstable empirical foundations for theory-building [44]. More broadly, linguistics has entered the wider "replicability" conversation. A special issue on the replication crisis in linguistics highlights that limited method-

ological transparency and shifting analytic choices can undermine reproducibility, especially when constructs and operationalisations vary across subfields [45]. One study argues that replication challenges in linguistics may also reflect genuine properties of language as a social phenomenon interacting with context and variation, which makes cross-method triangulation (corpus + experiment + field data) even more essential rather than optional [46]. Language acquisition relies on experiments with children; neurolinguistics uses imaging; NLP relies on statistical learning. These varied cultures set different standards of rigor and evidence, making cross-field collaboration difficult. Even terminology differs across fields, introducing "terminological gaps" and confusion. For instance, the concept of "competence" in formal linguistics differs from "communicative proficiency" in education. This is also evidenced by disagreement over clinical terms in clinical and neurolinguistics, as different usage of terminology by experts can cause confusion with regards to the concrete concept mentioned. With clinical terminology, disagreement over definitions can have real adverse downstream costs for individuals seeking help from clinical practice [47].

Without deliberate coordination, these differences create blind spots and duplicated debates. The difference in terminology between fields is not only confusing for linguists but also comes with considerable concerns with regards to the nature of linguistics as a science. One review discusses the presence of pseudoscientific theories in linguistics and is able to identify several unfalsifiable hypotheses in different linguistic disciplines, which potentially arise in part from the nature of current linguistic inquiry. This relates to speaker performance yet purports to produce insights regarding speaker competence [42]. Since research on real-world language use primarily concerns performance, hypotheses about underlying speaker competence are difficult to test empirically, as the abstract nature of linguistic competence makes direct falsification challenging. Importantly, recent "data governance" work in NLP has tried to reduce this conceptual opacity by formalising what data actually represent: another paper argues that standardised "data statements" can make dataset provenance, population coverage, and collection context explicit, enabling better science and reducing bias [48]. This is a concrete example of a field attempting to close terminological and evidential gaps by institutionalising documentation norms.

A key flaw is epistemic bias, as the large psycholinguistic survey found, linguistic diversity is often glossed over; most research assumes Western, educated, industrialised languages and speakers. According to a systematic review, 76% of all studies concentrate on just ten languages, nearly half of which (45%) focus on English, risking theories that reflect language-specific rather than universal processing principles [49]. This skew not only limits our scientific conclusions, it also raises equity issues. Likewise, sign languages, creoles, and emergent digital communication are under-studied in many subfields. Research on multilingualism and heritage languages often proceeds with little interaction between psycholinguistics, sociolinguistics, and education. In documentation, the common assumption of monolingual fields obscures the fact that most communities are multilingual, as Holton notes. These implicit biases arise from siloed training: scholars tend to study what is familiar (their native tongue, their own society), which weaves systemic gaps into the fabric of language science. Parallel evidence from multilingual NLP shows the same structural problem at scale: dataset and model development strongly tracks resource availability rather than theoretical importance, producing systematic inequities in performance for low-resource languages and dialects, and reinforcing English-centric "default assumptions" in technology design [50,51]. The key structural point is that "bias" is not just a moral issue; it is an epistemic constraint produced by incentives, infrastructure, and training pipelines.

Community engagement is another structural weakness. Many language subfields, particularly anthropological linguistics and documentation champion community-based research. They emphasise informed consent, local leadership, and benefit sharing. By contrast, much research in linguistics and NLP happens with secondary data or in labs, with minimal direct involvement of language communities. This raises ethical concerns: who controls the data, who benefits from new technologies, and whose knowledge is prioritised? Without coordination, projects can violate community norms (e.g., recording sacred language without permission) or fail to return results. The Linguistic Society of America's ethics guidelines stress that language knowledge is collective and must be handled with respect. However, the current fragmented landscape of language research limits the consistent application of these best practices across subfields. In addition, existing institutional workflows are often not equipped to accommodate the requirements of post-publication accountability and open science. Uneven community engagement is a predictable outcome when subfields operate under incompatible infrastructure and reward systems.

5. Addressing Objections and Previous Attempts at Integration

As we highlight these problems, we must address common concerns. One worry is that an integrated language science might dilute the formal rigor of theoretical linguistics. In fact, a pluralistic language science can preserve and even strengthen rigor by applying formal methods across domains. Programs like ILST explicitly include "structural" linguistics alongside computational and cognitive approaches [9]. The Stanford Encyclopedia of Philosophy notes that computational linguistics has long drawn on formal grammar, logic, and cognitive models. In a unified framework, formal theory can guide AI models (e.g., providing constraints for learning algorithms), and quantitative methods can be applied to test linguistic theories with real-world data. Thus, retiring linguistics as a silo does not mean abandoning grammar or semantics, it means embedding them in a richer empirical context.

At the same time, critics may argue that "formal rigor" is not just about having formal tools available somewhere in the ecosystem, but about what gets rewarded and reproduced in training and publishing. A common fear is that integration shifts incentives away from explanatory adequacy (clear assumptions, explicit derivations, and theory comparison) toward engineering-style optimisation (leaderboards, benchmark chasing, and aggregate scores), which can make it harder for formally precise work to remain central rather than peripheral. This concern is not hypothetical in computational linguistics and NLP: the field has repeatedly debated whether performance gains translate into understanding, and whether evaluation regimes push researchers toward opaque scaling rather than interpretable, theory-linked modelling. For example, work on "human-scale" language modelling argues that the current scaling paradigm brings scientific downsides as models trained on trillions of words may excel at tasks while becoming less informative as models of human language learning and processing [52].

Another objection is fear of duplication or loss of identity; if "everyone studies language," won't we see redundant research? The solution lies in collaboration: a community of language scientists can coordinate projects so that efforts are complementary rather than overlapping. For example, when Penn created an interdisciplinary unit spanning cognitive science and linguistics, it reduced duplicate courses and unified resources. Grant agencies could similarly encourage integrated proposals that pool expertise. The desired outcome is not identical replication, but synergy, leveraging diverse strengths. For instance, an AI research team and a field linguist might jointly collect a speech corpus annotated with syntactic structure, benefiting both language model training and typological analysis. Likewise, the implementation of computational methodologies and corpora can in turn assist formal syntax by informing grammaticality judgements and reducing the need for introspection.

Sceptics may say that linguistics already implies multidisciplinarity in practice (phonetics drawing on physiology, psycholinguistics on cognitive psychology, sociolinguistics on sociology/anthropology, computational linguistics on logic and CS), and that "integration" risks becoming a rebranding exercise rather than a substantive change in how questions are formulated. On this view, the real danger is not duplication of topics (which already happens), but duplication of incompatible standards; different subfields may answer "the same" question using different ontologies and integration could intensify confusion unless it also forces explicit articulation of what counts as evidence and explanation. Interdisciplinary collaborations often involve negotiation of epistemic identities; what methods are legitimate, what counts as a good explanation, and who has authority to define the problem. In other words, critics may worry that an integrated language science could increase friction and conceptual mismatch even while claiming unity unless it builds robust shared frameworks rather than simply collocating researchers.

Critiques of modern NLP also illustrate the need for integration. It has been argued that current neural models are essentially "curve-fitting" on form, lacking true semantic understanding as they capture statistical regularities rather than meaning, often misinterpreting nuance, figurative language, or pragmatic intent. Many models also rely on simplified assumptions such as conditional independence or shallow contextualisation, which fail to reflect the structural and relational complexity of natural language [53]. In an integrated language science NLP would routinely incorporate insights from psycholinguistics, semantics, and social context. An article on natural language understanding argues that form-only systems cannot learn meaning without grounding [54]. A unified field would address this by connecting large-scale data methods with rigorous semantic theory and human behavioral evidence. Likewise, concerns that AI-based language tools are insensitive to user context can be met by incorporating sociolinguistic and cognitive factors into design. In short, objections about rigor or redundancy would spur careful structuring of research programs to ensure that formal insights are preserved, overlap is managed, and meaning is not lost in abstraction.

6. Toward an Integrated Language Science: Institutional and Epistemic Reforms

What would a fully integrated Language Science look like? Conceptually it would treat language research as a continuum across levels—from neurons to societies—rather than as separate disciplines. Practically, universities might establish transdisciplinary Language Science centres or institutes (as some already have), with faculty appointments that span departments. Graduate programs would train students in computational modelling, experimental methods, field documentation, and societal aspects together, so that a PhD might be co-advised by a neuroscientist and a sociolinguist for example. Course curricula would blend theory and application, an introductory "language science" course could cover generative grammar and neural networks, acquisition experiments and corpus analysis, all as facets of the same phenomenon. The mission statement of the Language Science Center at UMD captures this vision: its interdisciplinary approach "addresses broad questions and combines expertise in multiple departments, methodologies, and research areas".

To make this continuum operational (and not just rhetorical), an integrated field would explicitly map "levels" to linked kinds of evidence and linked outputs. For instance, the same core question, how humans represent structure and meaning, could be pursued via (i) formal grammatical constraints and semantic composition (theory level), (ii) behavioral signatures like reaction times or eye movements (cognition level), (iii) neural dynamics measured with EEG/fMRI/intracranial recordings (brain level), and (iv) large-scale patterns in corpora that capture typology, register, and social variation (population level). The point being that not every project must cover all levels, but that the field would build "interfaces" so findings can travel across levels: formal predictions become testable hypotheses in experiments; experimental patterns inform model constraints; corpus and sociolinguistic variation define what counts as realistic input distributions for computational models; and computational models generate new testable predictions about cognition or change. A concrete example of this "interface building" is the rise of computational sociolinguistics, which explicitly positions itself as a bridge between sociolinguistic theory and large-scale NLP methods rather than as a separate niche [55].

In an institutional sense, a comprehensive Language Science model would also be team-science by design. That means centres would provide shared project management structures, transparent leadership, and mechanisms for mutual learning so that cross-field teams can actually function (and not collapse into parallel play). Modern team-science work on interdisciplinary collaboration emphasises that successful cross-disciplinary projects require explicit shared goals, distributed leadership, and social learning processes [56].

In terms of outputs, new journals and conferences should welcome cross-field work. For example, the Frontiers journal launched a "Computational Sociolinguistics" topic precisely to unite social linguistics and NLP [56]. We envision more such venues. Similarly, the success of Language Science Press (an open-access publisher)—277 books in multiple languages as of 2025—shows that the linguistics community values broad, cross-cutting scholarship [55].

In a fully integrated Language Science ecosystem, these venues would not just "accept interdisciplinary work," but would normalise hybrid contribution types, for example: theory papers that ship with reusable datasets or annotation schemes; computational papers that include explicit links to linguistic constructs and error analyses grounded in typology or pragmatics; experimental papers that preregister predictions derived from formal or computational models; resource papers that treat corpora, lexicons, and benchmarks as first-class scientific contributions (especially for under-resourced languages and varieties). This matters because, as recent sociolinguistic perspectives on language modelling argue, language models are inherently models of varieties of language, so corpus design and representativeness become theoretically consequential [57].

A key piece of the "comprehensive model" is shared infrastructure. A mature Language Science centre would typically maintain certain components. First, a shared data core (corpora, audio/video repositories, experimental stimuli, and metadata standards) that supports both computational work and field documentation, with clear licensing and community consent where relevant. Next, a methods core (labs for psycholinguistics, neuro methods, and computational resources) plus training support so students can actually move between toolchains. Importantly, it would have a theory-measurement interface (working groups where formal linguists, experimentalists, and modelers align on constructs: e.g., what counts as "agreement," "scope," "reference," "information structure," etc., and how those constructs will be operationalised in tasks and annotations). Finally, a translation and impact core (clinical language assessment ties, education partnerships, speech-tech/NLP evaluation in real settings, policy input on language rights, and public communication) would be needed for successful functioning.

Funding bodies should offer "language science" grants that require multi-domain teams, and professional societies (Linguistic Society of America (LSA), Association for Computational Linguistics (ACL), Society for Neuroscience (SfN), American Anthropological Association (AAA), etc.) could form a coalition for language research. Governance of language science might involve cross-divisional committees (e.g., joint sessions at major science funding agencies) to reflect that language belongs in humanities, social science, natural science, and engineering alike. In practice, that coalition could be organised around recurring "grand challenge" programs that force integration without forcing uniformity, for example: (i) Language acquisition across diverse inputs, (ii) Language and social variation at scale, (iii) Neural and computational mechanisms of composition, (iv) Language breakdown and recovery, (v) Meaning, grounding, and interaction, and (vi) Multilingual equity and documentation. Each program would have clear roles for formal theory, experiments, corpora, and modelling, with shared evaluation plans and shared deliverables. The reason to foreground multilingual equity is scientific: cognitive-science and language-generalisation problems are worsened by overreliance on English, and recent work argues that English-centric sampling can distort claims about "how language works" more generally.

Finally, an integrated Language Science would make two things routine that are currently uneven across subfields: (1) explicit documentation of "what language" is being studied (which population, which variety/register, what communicative context), and (2) explicit alignment between claims and evidence (e.g., a model paper is clear whether it is a tool, a cognitive hypothesis, or a theory of linguistic structure). This is exactly the kind of friction-reducing integration that computational sociolinguistics discussions point to: collaboration works best when the methodological power of NLP is paired with sociolinguistic theory and careful attention to bias, replicability, and representativeness, especially as LLMs become easier to use and more tempting to treat as generic language engines [58].

7. Evidence of Cross-Silo Success

There are encouraging examples where integration has yielded new insights. In language documentation and revitalisation, collaborations between communities, linguists, and technologists have led to practical tools. For instance, several companies (e.g., Google, Microsoft) now partner with linguists to develop NLP for endangered languages. A recent overview notes that machine-learning speech recognition and translation pipelines for low-resource languages, developed in partnership with academic centres, can accelerate documentation and learning resources. Concrete examples include Google's collaboration with the Centre of Excellence for the Dynamics of Language, which produced automatic speech recognition pipelines for languages with very small speaker populations, dramatically reducing the time required to transcribe and analyse oral language data [59]. These efforts show that when linguistic fieldwork, sociolinguistic expertise, and machine learning are integrated, the resulting tools not only advance research but also directly support community-driven revitalisation by producing usable dictionaries, transcripts, and educational materials.

The integration of linguistics and neuroscience has likewise led to major breakthroughs in understanding and applying knowledge about language in the brain. A landmark example is the study by Huth et al. (2016), which used fMRI and voxel-wise modelling to map semantic representations across the cerebral cortex while participants listened to natural stories [60]. Their results revealed large-scale, consistent semantic maps spanning temporal, parietal, and frontal regions, demonstrating that meaning is represented in distributed cortical systems rather than localised symbolic modules. Subsequent work has extended this approach by combining naturalistic stimuli with representations derived from large language models, showing that computational semantic embeddings grounded in linguistics and NLP can predict neural responses across individuals and modalities [11]. At the interface of linguistics, neurosurgery, and machine learning, research demonstrates that speech could be decoded directly from cortical activity and synthesised into intelligible spoken output using a brain-computer interface informed by articulatory phonetics [61]. By explicitly modelling the vocal tract and speech acoustics, this system generated natural-sounding sentences that listeners could understand, illustrating how linguistic theory can directly guide neural decoding to restore communication in people with severe speech impairments.

Clinical and applied domains further demonstrate the payoff of cross-silo integration. In aphasia therapy, approaches such as Intensive Language Action Therapy combine linguistic theory, neurorehabilitation, and psycholinguistic assessment to improve outcomes for stroke patients, even years after injury. Randomised controlled trials and longitudinal studies show that intensive, language-focused interaction can induce durable improvements in

chronic aphasia, accompanied by measurable changes in language-related brain networks [62].

More recently, linguistics has become central to computational psychiatry and clinical assessment, where NLP analyses of spontaneous speech, drawing on discourse coherence, referential cohesion, and pragmatic structure, function as scalable digital phenotypes for psychosis risk, symptom tracking, and disease progression [63]. Importantly, theoretical work emphasises that these linguistic markers are biosocial rather than purely biological, requiring explicit integration of sociolinguistic context to avoid reductionist clinical interpretations [64]. In neurodegenerative disease, interdisciplinary teams spanning neurology, geriatrics, speech-language pathology, and computational linguistics have shown that connected-speech analysis can detect cognitive decline and Alzheimer's-related impairment from short, naturalistic samples, positioning linguistic structure and prosody as clinically meaningful, low-burden screening tools [65].

Computational sociolinguistics has emerged as a hybrid field in which researchers use large-scale social data and network analysis to study language variation and change while applying sociolinguistic theory to improve NLP models. Survey work documents how this field explicitly bridges sociolinguistic constructs such as identity, style, and register with computational methods operating at population scale [66]. The rapid expansion of computational sociolinguistics over the past decade shows that integration enables questions about language variation, social meaning, and change to be addressed at unprecedented scale by teams combining linguistic, social-scientific, and data-science expertise. Taken together, these cases demonstrate that when linguistics is embedded within genuinely interdisciplinary research programs, it generates empirical findings, scalable applications, and theoretical advances that could not have emerged from any single disciplinary tradition alone.

8. Ethical and Equity Considerations

A reinvigorated language science must also foreground ethics and equity. Researchers should adopt a community-based participatory approach: linguists and data scientists must engage language communities from project inception. This means co-designing research goals, ensuring informed consent (which may be an ongoing dialogue, not a one-time form), and agreeing on who will access and steward the resulting data. Linguists have long warned that language knowledge is "cultural and collective" and must be handled with respect. In practice, this entails negotiating with local leaders for consent (especially where community norms give chiefs a role), documenting permissions in whatever form is appropriate (written, recorded, or verbal), and allowing participants to withdraw at any time.

Recent ethics scholarship across participatory research underscores why "from inception" matters: when communities are involved only at the data-collection stage, key decisions about framing, risk, representation, and downstream uses have already been made, often in ways that communities would not endorse. Shared governance should be maintained throughout the entire research process, from design and data collection to analysis, publication, and beyond. Ethical issues often arise after data collection, including how findings represent the community, how authorship is determined, and how materials are used outside their original context. For language science specifically, this lifecycle perspective becomes more urgent when language documentation is linked to NLP and AI. Data that appear low-risk as recordings or transcripts can become high-stakes once aggregated, searchable, and reused in new technical contexts, such as training speech recognition systems or large language models. This has motivated a shift away from purely individual consent models toward collective data governance. The CARE Principles for Indigenous Data Governance: Collective Benefit, Authority to Control, Responsibility, and Ethics, were developed to complement open science norms by explicitly recognising community-level rights over cultural and linguistic data [67].

Benefit-sharing is crucial. Ethical documentation practice urges compensating speakers for their time and expertise. This compensation should be culturally appropriate, monetary or otherwise, and should not create coercion or jealousy. For instance, researchers might fund community-driven educational materials or infrastructure instead of paying per hour of recording. Crucially, language communities should gain access to research outputs and tools (e.g., dictionaries, learning apps) on equitable terms. Technology developers can partner with communities to tailor NLP tools (speech-to-text, translation) to local needs, thus sharing profits or services. Transparency about data use is also ethical: communities must know if their recordings will be used to train a for-profit model or published openly.

In the last decade, benefit-sharing has been operationalised through concrete practices such as shared gov-

ernance of datasets, inclusive authorship models, and community capacity-building. Participatory NLP research demonstrates that so-called "low-resource" languages are not naturally data-poor but structurally marginalised. Work on participatory machine translation for African languages argues that ethical NLP requires long-term collaboration, local expertise, and infrastructural investment rather than extractive data collection [68].

Transparency about data use is increasingly supported by standardised documentation practices. In NLP, "data statements" have been proposed to explicitly document who the data represent, how they were collected, consent conditions, and limitations on reuse. These statements are designed to prevent inappropriate generalisation and to surface ethical risks before models are trained or deployed, directly supporting community decision-making about data sharing and downstream applications [48]. As language technologies scale, transparency must also extend to downstream harms. Recent risk-mapping work on large language models identifies systematic risks including representational harm, exclusion, discrimination, and misuse. These frameworks are increasingly used to inform ethical review, consent negotiations, and governance decisions, especially when community language data are used in commercial or high-impact systems [69]. Equity demands that integrated language science actively expand its focus beyond the English-centric and Global North bias. Funding and publication policies should prioritise research on under-studied languages and meaningfully involve scholars from the Global South. As work on sampling bias in applied linguistics shows, over-reliance on WEIRD populations distorts empirical conclusions and shapes citation practices, reinforcing epistemic inequality [70]. This critique is strengthened by recent quantitative audits demonstrating that psycholinguistic and cognitive language research remains heavily concentrated in a small number of countries and languages, despite increasingly global claims. Such patterns limit theory-building and obscure language diversity, even as they disproportionately disadvantage scholars and communities outside dominant research centres [49].

Finally, equity is not only about who is studied but also about who can participate as researchers. Recent work on Indigenous data sovereignty and cultural heritage governance emphasises the need for institutional infrastructures, such as ethical templates, multilingual consent procedures, community partnership agreements, and flexible data-access regimes, that enable equitable collaboration rather than placing the full ethical burden on individual researchers [71]. In an integrated Language Science, these ethical and equity commitments are not peripheral concerns but structural requirements. Building systems that center community agency, transparency, and global inclusion is essential for scientific validity: understanding human language in its full complexity requires ethical partnerships with the people who speak it.

9. Conclusions

Retiring linguistics as a silo is an ambitious proposal, but one that follows directly from the realities of contemporary language research. Language is already studied across neuroscience, psychology, computation, education, clinical practice, and social life, yet institutional structures continue to organise these domains as if they were separable. The goal is not to reject linguistics, but to carry its foundational contributions: formal theory, descriptive precision, and ethical awareness, into a framework that reflects how language is actually investigated today. Across fields, calls for integration have emerged not as speculative ideals but as responses to persistent structural fragmentation. Without such reform, the organisation of language research will remain misaligned with both empirical practice and the accelerating influence of technology on communication.

Realising this vision requires structural change rather than rhetorical commitment. Universities and funding agencies must establish and sustain language-science institutes that cut across traditional boundaries. Evaluation and tenure systems must treat collaborative, interdisciplinary work as central scholarly labour, not as an exception to disciplinary norms. Curricula must present language in its full complexity, integrating formal theory with field methods, quantitative analysis with cultural and social context. Journals and conferences must treat multilevel research as standard rather than peripheral, and where necessary create venues explicitly designed for such work. The research community must also recognize and support scholars whose expertise spans approaches, as this intellectual breadth is essential rather than ancillary.

An integrated Language Science enables questions to be addressed at the scale they demand, drawing simultaneously on neural, cognitive, social, and cultural evidence. In the presence of ongoing fragmentation, such integration is not optional: it is the only framework capable of treating language as a single, multi-level system rather than as a collection of disconnected problems. The consequences of integration are both scientific and societal:

stronger theory, more effective language education, more reliable and equitable language technologies, improved clinical interventions for communication disorders, and better-informed policies for multilingual communities. As language becomes increasingly central to technology, health, and governance, maintaining disciplinary silos is no longer defensible. Retiring linguistics, in this sense, means retiring its isolation, and replacing it with a Language Science commensurate with the complexity and importance of its subject matter.

Author Contributions

Conceptualisation, I.A. and T.E.V.; methodology, I.A.; original draft preparation, I.A. and T.E.V.; review and editing, I.A. and T.E.V. Both authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

This study does not involve human participants, animal subjects, or primary data collection. Therefore, ethical approval was not required.

Informed Consent Statement

Not applicable.

Data Availability Statement

This study is a theoretical and conceptual analysis and does not involve primary data collection. No datasets were generated or analysed during the current study. All referenced materials are publicly available through the cited sources.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

1. Chomsky, N. Persistent topics in linguistic theory. *Diogenes* **1965**, *13*, 13–20. [CrossRef]
2. Pinker, S. *The Language Instinct: The New Science of Language and Mind*; Penguin Books: London, UK, 1994.
3. Evans, N.; Levinson, S.C. The myth of language universals: Language diversity and its importance for cognitive science. *Behav. Brain Sci.* **2009**, *32*, 429–448. [CrossRef]
4. Bybee, J.L. *Language Change*; Cambridge University Press: Cambridge, UK, 2015.
5. Lieven, E. Usage-based approaches to language development: Where do we go from here? *Lang. Cogn.* **2016**, *8*, 346–368. [CrossRef]
6. Goldberg, A.E. *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*; Princeton University Press: Princeton, NJ, USA, 2019.
7. Tomasello, M. The ontogeny of cultural learning. *Curr. Opin. Psychol.* **2016**, *8*, 1–4. [CrossRef]
8. Levy, R.P. Communicative efficiency, uniform information density, and the rational speech act theory. In Proceedings of the Annual Meeting of the Cognitive Science Society, Madison, WI, USA, 25–28 July 2018.
9. ILST. Available online: https://web.sas.upenn.edu/langscience/ (accessed on 27 May 2025).
10. Willems, R.M.; Frank, S.L.; Nijhof, A.D.; et al. Prediction during natural language comprehension. *Cereb. Cortex* **2015**, *26*, 2506–2516. [CrossRef]
11. Goldstein, A.; Zada, Z.; Buchnik, E.; et al. Shared computational principles for language processing in humans and deep language models. *Nat. Neurosci.* **2022**, *25*, 369–380. [CrossRef]

12. Fedorenko, E.; Ivanova, A.A.; Regev, T.I. The language network as a natural kind within the broader landscape of the human brain. *Nat. Rev. Neurosci.* **2024**, *25*, 289–312. [CrossRef]
13. Diachek, E.; Blank, I.; Siegelman, M.; et al. The domain-general multiple demand (MD) network does not support core aspects of language comprehension: A large-scale fMRI investigation. *J. Neurosci.* **2020**, *40*, 4536–4550. [CrossRef]
14. Alday, P.M. M/EEG analysis of naturalistic stories: A review from speech to language processing. *Lang. Cogn. Neurosci.* **2018**, *34*, 457–473. [CrossRef]
15. Zhao, J.; Gao, R.; Brennan, J.R. Decoding the neural dynamics of headed syntactic structure building. *J. Neurosci.* **2025**, *45*, e2126242025. [CrossRef]
16. Christiansen, M.H.; Chater, N. The Now-or-Never bottleneck: A fundamental constraint on language. *Behav. Brain Sci.* **2015**, *39*, e62. [CrossRef]
17. Hasson, U.; Egidi, G.; Marelli, M.; et al. Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension. *Cognition* **2018**, *180*, 135–157. [CrossRef]
18. Kirby, M.S.; Spencer, T.D.; Spiker, S.T. Humble behaviorism redux. *Behav. Soc. Issues* **2022**, *31*, 133–158. [CrossRef]
19. Arnon, I.; Kirby, S. Cultural evolution creates the statistical structure of language. *Sci. Rep.* **2024**, *14*, 5255. [CrossRef]
20. Futrell, R. Information-theoretic principles in incremental language production. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120*, e2220593120. [CrossRef]
21. Koplenig, A. Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *R. Soc. Open Sci.* **2019**, *6*, 181274. [CrossRef]
22. Henrich, J.; Heine, S.J.; Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **2010**, *33*, 61–83. [CrossRef]
23. Kidd, E.; Garcia, R. How diverse is child language acquisition research? *First Lang.* **2022**, *42*, 014272372110664. [CrossRef]
24. Chomsky, N. Derivation by phase. In *A Life in Language*; MIT Press: Cambridge, MA, USA, 2001; pp. 1–52. [CrossRef]
25. Keine, S.; Zeijlstra, H. Clause-internal successive cyclicity: Phasality or DP intervention? *Nat. Lang. Linguist. Theory* **2025**, *43*, 1119–1182. [CrossRef]
26. Blümel, A.; Collins, C. Smuggling and labeling theory. *Biolinguistics* **2025**, *19*. [CrossRef]
27. Wang, N.; Li, J. Estimating the strength and timing of syntactic structure building in naturalistic reading. *arXiv preprint* **2025**, *arXiv:2509.23195*. [CrossRef]
28. Arehalli, S.; Dillon, B.; Linzen, T. Syntactic surprisal from neural models predicts, but underestimates, human processing difficulty from syntactic ambiguities. *arXiv preprint* **2022**, *arXiv:2210.12187*. [CrossRef]
29. Smith, N.J.; Levy, R. The effect of word predictability on reading time is logarithmic. *Cognition* **2013**, *128*, 302–319. [CrossRef]
30. Friederici, A.D. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* **2002**, *6*, 78–84. [CrossRef]
31. Harris, D.W.; Rubio-Fernandez, P. Common ground: Between formal pragmatics and psycholinguistics. Available online: https://philarchive.org/rec/HARCGF (accessed on 27 May 2025).
32. Scontras, G.; Degen, J.; Goodman, N.D. On the grammatical source of adjective ordering preferences. *Semant. Pragmat.* **2019**, *12*, 1–21. [CrossRef]
33. Port, R.F.; O'Dell, M. Neutralization of syllable-final voicing in German. *J. Phonetics* **1985**, *13*, 455–471. [CrossRef]
34. Du, N.; Durvasula, K. Phonetically incomplete neutralisation can be phonologically complete: Evidence from Huai'an Mandarin. *Phonology* **2022**, *39*, 559–595. [CrossRef]
35. Merrill, W.; Goldberg, Y.; Schwartz, R.; et al. Provable limitations of acquiring meaning from ungrounded form: What will future language models understand? *Trans. Assoc. Comput. Linguist.* **2021**, *9*, 1047–1060. [CrossRef]
36. Schwarz-Friesel, M. On the status of external evidence in the theories of cognitive linguistics: Compatibility problems or signs of stagnation in the field? Or: Why do some linguists behave like Fodor's input systems? *Lang. Sci.* **2012**, *34*, 656–664. [CrossRef]
37. Priem, J.; Piwowar, H.; Orr, R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint* **2022**, *arXiv:2205.01833*. [CrossRef]

38. Oxley, K. Partial and particularistic: Grant peer review panels evaluating interdisciplinarity. *Res. Eval.* **2024**, *34*, rvaf037. [CrossRef]
39. Xiang, S.; Romero, D.M.; Teplitskiy, M. Evaluating interdisciplinary research: Disparate outcomes for topic and knowledge base. *Proc. Natl. Acad. Sci. U.S.A.* **2025**, *122*. [CrossRef]
40. Rushforth, A.; De Rijcke, S. Practicing responsible research assessment: Qualitative study of faculty hiring, promotion, and tenure assessments in the United States. *Res. Eval.* **2024**, *33*, rvae007. [CrossRef]
41. Mäkinen, E.I.; Evans, E.D.; McFarland, D.A. Interdisciplinary research, tenure review, and guardians of the disciplinary order. *J. High. Educ.* **2024**, *96*, 54–81. [CrossRef]
42. Eddington, D. Linguistics and the scientific method. Available online: https://www.researchgate.net/publication/228848641_Linguistics_and_the_scientific_method (accessed on 27 May 2025).
43. Gibbs, R.W. Introspection and cognitive linguistics: Should we trust our own intuitions? *Annu. Rev. Cogn. Linguist.* **2006**, *4*, 135–151. [CrossRef]
44. Linzen, T.; Oseki, Y. The reliability of acceptability judgments across languages. *Glossa* **2018**, *3*. [CrossRef]
45. Sönning, L.; Werner, V. The replication crisis, scientific revolutions, and linguistics. *Linguistics* **2021**, *59*, 1179–1206. [CrossRef]
46. Grieve, J. Observation, experimentation, and replication in linguistics. *Linguistics* **2021**, *59*, 1343–1356. [CrossRef]
47. Bishop, D.V.M. Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *Int. J. Lang. Commun. Disord.* **2017**, *52*, 671–680. [CrossRef]
48. Bender, E.M.; Friedman, B. Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Trans. Assoc. Comput. Linguist.* **2018**, *6*, 587–604. [CrossRef]
49. Berghoff, R.; Bylund, E. Diversity in research on the psychology of language: A large-scale examination of sampling bias. *Cognition* **2025**, *256*, 106043. [CrossRef]
50. Zhu, S.; Supryadi; Xu, S.; et al. Multilingual large language models: A systematic survey. *arXiv preprint* **2024**, *arXiv:2411.11072*. [CrossRef]
51. Joshi, A.; Dabre, R.; Kanojia, D.; et al. Natural language processing for dialects of a language: A survey. *ACM Comput. Surv.* **2025**, *57*, 149. [CrossRef]
52. Wilcox, E.G.; Hu, M.Y.; Mueller, A.; et al. Bigger is not always better: The importance of human-scale language modeling for psycholinguistics. *J. Mem. Lang.* **2025**, *144*, 104650. [CrossRef]
53. Su, F. Research on integration of emotion analysis in English modular teaching based on natural language processing. *Front. Psychol.* **2022**, *13*. [CrossRef]
54. Bender, E.M.; Koller, A. Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 5185–5198. [CrossRef]
55. Nguyen, D.; Doğruöz, A.S.; Rosé, C.P.; et al. Computational sociolinguistics: A survey. *Comput. Linguist.* **2016**, *42*, 537–593. [CrossRef]
56. Gesing, P.; Tornwall, J.; Kulo, V.; et al. Team science in interdisciplinary health professions education research: A multi-institutional case study. *Adv. Health Sci. Educ.* **2025**, *30*, 1123–1141. [CrossRef]
57. Language Science Press. Available online: https://langsci-press.org (accessed on 27 May 2025).
58. Grieve, J.; Bartl, S.; Fuoli, M.; et al. The sociolinguistic foundations of language modeling. *Front. Artif. Intell.* **2025**, *7*. [CrossRef]
59. Blasi, D.E.; Henrich, J.; Adamou, E.; et al. Over-reliance on English hinders cognitive science. *Trends Cogn. Sci.* **2022**, *26*. [CrossRef]
60. Huth, A.G.; de Heer, W.A.; Griffiths, T.L.; et al. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **2016**, *532*, 453–458. [CrossRef]
61. Anumanchipalli, G.K.; Chartier, J.; Chang, E.F. Speech synthesis from neural decoding of spoken sentences. *Nature* **2019**, *568*, 493–498. [CrossRef]
62. Dreyer, F.R.; Doppelbauer, L.; Büscher, V.; et al. Increased recruitment of domain-general neural networks in language processing following intensive language-action therapy: fMRI evidence from people with chronic aphasia. *Am. J. Speech Lang. Pathol.* **2021**, *30*, 455–465. [CrossRef]
63. Morgan, S.E.; Diederen, K.; Vértes, P.E.; et al. Natural language processing markers in first episode psychosis and people at clinical high-risk. *Transl. Psychiatry* **2021**, *11*, 630. [CrossRef]
64. Palaniyappan, L. More than a biomarker: Could language be a biosocial marker of psychosis? *NPJ Schizophr.* **2021**, *7*, 42. [CrossRef]
65. Balabin, H.; Tamm, B.; Spruyt, L.; et al. Natural language processing-based classification of early Alzheimer's

disease from connected speech. *Alzheimers Dement.* **2025**, *21*, e14530. [CrossRef]
66. Barbieri, C.; Blasi, D.E.; Arango-Isaza, E.; et al. A global analysis of matches and mismatches between human genetic and linguistic histories. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, e2122084119. [CrossRef]
67. Carroll, S.R.; Garba, I.; Figueroa-Rodríguez, O.L.; et al. The CARE Principles for Indigenous Data Governance. *Data Sci. J.* **2020**, *19*. [CrossRef]
68. Nekoto, W.; Marivate, V.; Matsila, T.; et al. Participatory research for low-resourced machine translation: A case study in African languages. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2020, Online, 16–20 November 2020; pp. 2144–2160. [CrossRef]
69. Weidinger, L.; Uesato, J.; Rauh, M.; et al. Taxonomy of risks posed by language models. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, 21–24 June 2022; pp. 214–229. [CrossRef]
70. Andringa, S.; Godfroid, A. Sampling bias and the problem of generalizability in applied linguistics. *Annu. Rev. Appl. Linguist.* **2020**, *40*, 134–142. [CrossRef]
71. Spano, I.; Zhang, Y. Indigenous data sovereignty in intangible cultural heritage governance: A complementary approach to public–private partnerships. *Int. J. Cult. Prop.* **2025**, *32*, 167–193. [CrossRef]