UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

# When there's more than one elephant in the room

Milin, Petar; Divjak, Dagmar

*DOI:*
10.53841/bpscog.2024.1.9.7

*License:*
None: All rights reserved

*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Milin, P & Divjak, D 2024, 'When there's more than one elephant in the room: Interdisciplinary approaches to the study of language', *Cognitive Psychology Bulletin*, vol. 1, no. 9, pp. 7-15.
https://doi.org/10.53841/bpscog.2024.1.9.7

Link to publication on Research at Birmingham portal

**Publisher Rights Statement:**
This is a pre-publication version of the following article: Milin, P. and Divjak, D. (2024) 'When there's more than one elephant in the room: Interdisciplinary approaches to the study of language', The Cognitive Psychology Bulletin, 1(9), pp. 7–15. doi:10.53841/bpscog.2024.1.9.7.

**General rights**
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

**Take down policy**
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 16. May. 2024

**Cognitive Psychology & Developmental Psychology Section Joint Conference 2023**
**Conference keynote**

# When there's more than one elephant in the room: Interdisciplinary approaches to the study of language

Petar Milin[a] and Dagmar Divjak[a,b]

[a] *Department of Modern Languages, University of Birmingham, Birmingham, United Kingdom*

[b] *Department of English Language & Linguistics, University of Birmingham, Birmingham, United Kingdom*

Email: ooominds@ooominds.org
Department of Modern Languages, Edgbaston, Ashley Building, University of Birmingham, B15 2TT, Birmingham, United Kingdom

Petar Milin, p.milin@bham.ac.uk – corresponding author
Web: https://www.birmingham.ac.uk/petarmilin; https://outofourminds.bham.ac.uk
ORCID: https://orcid.org/0000-0001-9708-7031
X: @petarmilin

Dagmar Divjak, d.divjak@bham.ac.uk
Web: https://www.birmingham.ac.uk/dagmardivjak; https://outofourminds.bham.ac.uk
ORCID: https://orcid.org/0000-0001-6825-8508
X: @dagmardivjak

**Abstract:** This concise overview paper introduces the work done by the Out Of Our Minds research group at the University of Birmingham, highlighting the need for organic interdisciplinarity in contemporary language science. By summarizing two case studies, we underscore the need for compatible methodologies that address shared research objectives. Despite initial enthusiasm for a multidisciplinary approach to language in the 1950s, subsequent research efforts often remained confined within specific scientific traditions. Recently, however, we have witnessed a resurgence of these foundational ideas. Crucially, Out Of Our Minds embodies a paradigm shift where linguists leverage rigorous operationalizations to test key theoretical notions, while psychologists broaden their understanding of empirical phenomena with ecological relevance and purposefulness. By synergizing the strengths of both disciplines, we advance our understanding of the complex and dynamic system of human language.

**Keywords:** language usage; language learning; linguistic emergence; grammatical tense and aspect; long-term memory; interdisciplinary research

Language appears peculiarly challenging for scientific scrutiny, and that stands in stark contrast with our, for the most part, experience of speaking, writing, listening, and reading as effortless activities. The complex cognitive machinery responsible for our uniquely human communicative capacities is covert from scientific (and indeed any other) deliberations. At best, we are left with the possibility to engage with *hypothetical constructs* that appeal to us as probable mechanisms and/or structures responsible for the emergence of language phenomena. These constructs, in themselves, have more or less theoretical allure and more or less empirical support. Thus, the two largest elephants hiding in the chambers of language pertain to (a) the information language users extract and store, and (b) the cognitive mechanisms which enable them to do so.

Part of the answer to the first question seems to be given already: our knowledge about language is *circumstantial*. It is primarily based on indirect evidence: on language use. We build that knowledge through a meticulous analysis of exemplars found in language corpora, or we measure responses – as elicited behaviour or brain activity – to carefully chosen linguistic stimuli. Sadly, however, corpus exemplars and/or participant responses are mere proxies (indicators, not indicata), and we cannot directly testify which structures and functions are *sufficient* and *necessary* for generating said exemplars and responses. This is when we put our trust in a hypothesised construct. That seems inevitable. It is the nature of the beast. The largest of all elephants. We are those blind people in that story, and natural languages are, thus, like a snake, or a tree, or a wall. The greatest risk, however, lies in our tendency to take the hypothesis for the truth. The power of a testable explanation becomes the explanation itself. As if we, language researchers of all kinds, follow Dali's recommendation "to spread confusion, not eliminate it". In the Out Of Our Minds research group [https://outofourminds.bham.ac.uk], we believe that a way through can be found in the full commitment to interdisciplinarity, which presupposes mutually complementary theories, shared testable hypotheses, as well as compatibility of research methodologies.

**A bit of history**

Apart from several other disciplines that occasionally turn attention to human language, only three of them have a vested interest in the study of it. At the centre of this imagined research space, we find Linguistics. Psychology and Engineering (Applied Sciences) have their own traditions too. Most of the time they remain within their own disciplinary boundaries, showing little curiosity to learn what's happening in the "neighbourhood". This is in part because today's science, generally speaking, is highly specialised, making it increasingly more challenging to become an expert in even one domain, and even more so to understand and engage with others' terminological apparatus, research questions and methods.

It might come as a surprise that this was not always the case. As early as the 50s of the last century, there were interdisciplinary attempts, showing both a high level of mutual, cross-disciplinary respect, as well as a clarity in understanding that disciplines need one another for the true advancement of the knowledge about language. For example, Gardner (1954) suggested that efforts must be made along the three related axes, with "(1) the linguist's conception of language as a structure of systematically interrelated units, (2) the learning theorist's conception of language as a system of habits relating signs to behaviour, and (3) the information theorist's conception of language as a means of transmitting information" (p. x). Similarly, Maclay (1973) describes the early years of interdisciplinarity between Linguistics and Psychology as characterized by extremely good relations between the two, with "a common commitment to an operationalist philosophy of science, and a division of labour that prevented a number of potential difficulties from becoming overt […] linguists were assigned the 'states of messages,' while psychologists assumed responsibility for the 'states of communicators' and also, by default, 'the processes of encoding and decoding'" (pp. 570-571).

After the publication of his review of Skinner's (1957) book "Verbal Behavior", Chomsky (1959) expelled learning from the linguistic scene. In addition to that, the Generativist Revolution also brought a stark demarcation of the research spaces: linguists should study language *competence*, and psychologists language *performance* (cf., Greene, 1972). In effect, Chomsky demanded a division of domains rather than a division of labour; as if Chomsky's take on modular cognition spilt over to the research culture itself, allowing for very little, if any at all, cross-fertilization between related disciplines. He was not particularly supportive of the applied sciences either, given the fact that he denied any possibility for a probabilistic approach, which would presumably appeal to engineers: "the notion 'probability of a sentence' is an entirely useless one, under any known interpretation of this term" (Chomsky, 1969, p. 57; also see Chomsky, 1957).

After decades of separation, once more we witness an increasing appetite for interdisciplinary work. In the previous period of disjoined efforts, psychologists and engineers would be granted permission to test certain linguistic notions (viz. Chomsky, 1969) in terms of, respectively, their cognitive plausibility in laboratory experiments or their usability in computational processing applications. It has become obvious, however, that we need to open a dialogue, and facilitate an exchange of ideas, rather than to school 'the others'. In our research group, to reiterate, we are committed to learning about complementary theories, devising shared testable hypotheses, and exploring possibilities to utilise a range of promising methodologies from across disciplines.

Through our daily practice of studying language, we are constantly reminded of how messy that domain of research truly is. At any imagined level of scrutiny, we observe the indivisible multidimensionality of language features, as well as those features' complex dynamics. If we assume for a second that words are easy enough to distinguish, we quickly find that even their simplest objectively determinable properties such as length and frequency of occurrence are mutually entangled. Furthermore, while a word's length typically remains constant over time, its frequency of use often changes. Words' meanings are even more whimsical, to the point that we can quite safely say that words do not carry any self-contained meaning, but that the meaning is "resolved" dynamically and contextually. This was pointed out by Harris (1954) and Firth (1957) decades ago; in fact, the famous Russian psychologist Lev Vygotsky came to that conclusion roughly two decades before Harris and Firth: "a word acquires its sense [smysl/смысл] from the context in which it appears; in different contexts, it changes its sense" (Vygotsky, 1934/2012, p. 305 of the Russian original, p. 259 of the English translation published in 2012).

**The emergence of structure from use**

The concept of *emergence* is the central hypothetical construct of usage-based linguistics, but it lacks an operational definition. Thus, we begin by proposing the fundamental cognitive process of *learning* as a plausible way of operationalising emergence: *if linguistic abstraction is emergent from usage, then it ought to be learnable* (Divjak & Milin, 2022). Importantly, we rely strictly on the most fundamental learning principles, such as Associative Learning (e.g., Pearce & Bouton, 2001; Enquist & Ghirlanda, 2005; Bouton, 2007), which we model computationally using Error-Correction rules as defined by Widrow-Hoff (1960), Rescorla-Wagner (1972), Sutton and Barto (1990) and others. This framework furthers our understanding of both linguistic emergence as well as language learning (Ellis, 2006, 2016; also see Ramscar & Yarlett, 2007 and Arnon & Ramscar, 2012).

The proposed operationalisation of emergence in terms of learning seems self-evident. Linguists are focused on understanding the fully emerged language system, which then leaves the process itself curiously absent from theoretical scrutiny. It would be, however, incorrect to jump to the conclusion that linguists are uninterested in the possibility of understanding the system's developed state by understanding its development. Julià (1983), for example, know that "[concentration] on 'what is learned' neglects the learning process itself" and that "[the] analysis of final products does little to suggest effective investigative strategy into the interaction between speakers and listeners, and the circumstances that bring their behaviour about, in short, the real data" (p. 92).

Psychologists show the opposite "bias": the inclination to focus on processes or functions, such as learning. Often, they leave aside the aim and the broader reasons for a given function; e.g., the purpose it may have for the biological system and/or for its evolutionary fitness. The possibility we find exciting and promising is that fully emerged linguistic abstractions can steer the process of learning by acting as a guiding purpose of learning. In a computational sense, such targets (also called *teachers*) are criteria that supervise the process of learning (for the Machine Learning sense of supervised

learning, see Haykin, 1999; Roy & Chakraborty, 2013). Conceptually, however, thinking about *what needs to be learned* constrains the formal account of learning further by providing a holistic frame and, thus, a degree of ecological validity. Learning is not only about the mechanism (i.e., the principle or the algorithm) but also about the entities and/or events in the environment whose relationships need to be learned, to represent the structure of learners' (animal, human, machine) world (cf. Rescorla, 1988, p. 152). At least in part, such considerations are related to Poggio's (2012) point about how biological learning machines evolved in the first place.

The two disciplines show complementary strengths to a remarkable extent, which is what those authors who discussed early interdisciplinary attempts must have had on their minds (viz. Gardner, 1954; Maclay, 1973) although Osgood(1968) rather referred to complementary "insufficiencies". In the following two sections, we will present two case studies that illustrates how multiples of "insufficiencies" can be turned into strengths.

**When the linguist needs the psychologist: to test the constructs' (cognitive) plausibility**

Linguists have a long-vested interest in understanding how time is expressed in language. Typically, the concepts of tense and aspect are used, yet both appear rather elusive (Binnick, 1991). Tense is generally accepted to refer to markers that express ways of locating events in time relative to the moment of speaking, such as English *-ed* and *-ing*. This appears simple enough, still the concept has been debated since antiquity. The standard definition of aspect is much more impenetrable: aspect would represent different ways of viewing the internal temporal constituency of an event. To obviate the utter vagueness of the notion of aspect, linguists further proposed a whole range of determining criteria, typically dichotomies such as telic/atelic, bounded/unbounded, foregrounded/back-grounded, none of which actually clarifies matters to the uninitiated.

English allows 12 tense-aspect combinations in total (tense: past, present, future; aspect: simple, perfect, progressive, perfect progressive), as illustrated in Table 1.

**Table 1.** Examples of English Tense (columns) and Aspect (rows) combinations.

| | **Past** | **Present** | **Future** |
|---|---|---|---|
| **Simple** | **Joe cooked.** | **Joe cooks.** | **Joe will cook.** |
| **Perfect** | **Joe had cooked.** | **Joe has cooked.** | **Joe will have cooked.** |
| **Progressive** | **Joe was cooking.** | **Joe is cooking.** | **Joe will be cooking.** |
| **Perfect progressive** | **Joe had been cooking.** | **Joe has been cooking.** | **Joe will have been cooking.** |

Yet not all 12 possible combinations that grammar tables theoretically allow for occur with all verbs in actual usage, let alone equally frequently. A usage-based linguistic approach that relies on language corpora helps us to *delineate the problem space* – providing empirical insights into what occurs and co-occurs and, hence, what would be a plausible set of learning outcomes (targets).

In our paper on the English tense-aspect system (Romain, Ez-zizi, Milin, & Divjak, 2022), we posed the following research questions: what tense-aspect combinations occur in usage, and what are the characteristics of learnability of these used combinations, given naturalistic input. We trained an error-driven associative learning model (Naïve Discrimination Learner, NDL: Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; also see Baayen, Chuang, Shafaei-Bajestan, & Blevins, 2019 for recent developments) starting from raw sentences extracted from the British National Corpus (BNC, Leech,

1992) to feed the algorithm with lexical input cues (i.e., the infinitive of the verb in question) or contextual input cues (other words or word n-grams in the sentence such as, e.g., past, few, years, few#years, few#sentences, past#few#years etc.), and tense-aspect combinations as the learning outcomes (e.g., past simple, present progressive etc.).

The end results of learning simulations revealed that usage of the system of English tense and aspect indeed revolves around two axes, but they are not tense and aspect. Instead, the key dimensions are simplex vs. complex cues, or lexical vs. contextual., as illustrated in Figure 1 where the pink bars represent lexical cues, and bars of other colours represent elements or combinations of elements other than the verb itself.

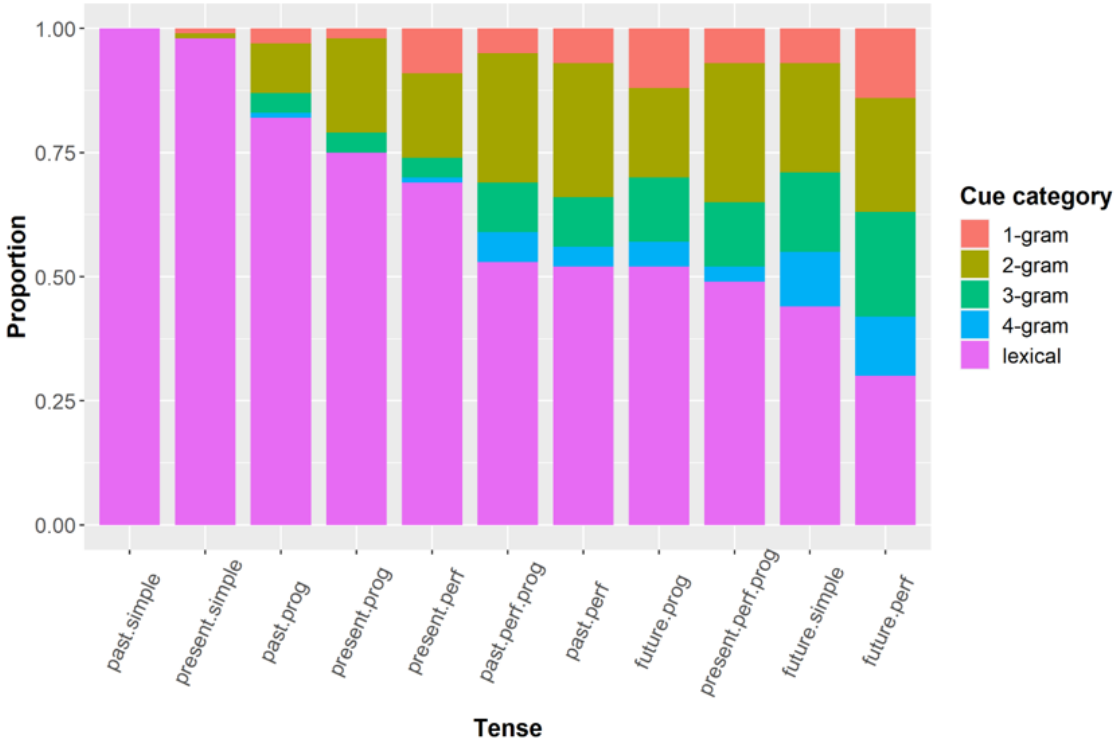


**Figure 1.** Distribution of the top 100 lexical and contextual cues for each TA combination (Romain et al., 2022).

Why do lexical cues come out so overwhelmingly strongly with the present and past simple? The present simple and past simple reflect temporal arrangements that are easy to conceptualize as the event takes place before, during or after the time of speaking. It is highly likely that this cognitive "simplicity" is the main reason why these two tenses are the most frequent in use. Enter learning, the high frequency of present past and past simple enables them to be learned from a vast range of competing lexical cues in usage experience. And because the number of times the TA is used exceeds the number of different lexical items, these items themselves become excellent cues for the TA. In other words, everything conspires to help us learn these simple tense-aspect combinations.

Conversely, however, the complex tense and aspect combinations confront us with a vicious circle. Arguably, such tense-aspect combinations require advanced event sequencing that reverse the order in which the events were experienced and are therefore limited to a specific type of narrative text. For example, while an eyewitness might say "*The officers **tried** to stem the flow of blood. The paramedics **arrived** and **took over***", using the simple past, the equivalent sentence retrieved from the British National Corpus reads "*Paramedics **were** at the scene in four minutes of the emergency call and **took***

***over*** *from officers who* ***had been trying*** *to stem the flow of blood*". Contexts requiring advanced event sequencing are cognitively more demanding and less often used. Due to the lower frequency of contexts requiring such complex TA forms, lexical items do not occur frequently enough to develop strong enough associations with these forms. Instead of building up strong lexical cues, complex TAs tend to be associated with contextual word n-grams, which are harder to spot and learn given the sheer number of such cues in usage. In other words, contrary to simple and lexically-cued tense-aspect combinations, everything conspires against learning the complex and contextually-cued tense-aspect combinations.

**When the psychologist needs the linguist: to propose a construct that needs testing**

The well-known divide between declarative and procedural (non-declarative, more recently) long-term memory in cognitive science (cf., Squire, Knowlton, & Musen, 1993; Squire, 2004, 2009), prompted research into how this memory division is implicated in processing and storing language. Notably, declarative and non-declarative long-term memory systems differ in storage and retrieval mechanisms. Declarative memory stores rapidly but demands conscious, controlled, and slow retrieval, while non-declarative memory, acquired slowly and unconsciously, exhibits automatic, reliable, and effortless retrieval. This division aligns well with the linguistic grammar-lexicon divide, where knowledge of syntax, morphology, and non-lexical semantics appear to be of a declarative nature, while the lexicon, being arbitrary and idiosyncratic, is of non-declarative nature (Ullman, 2004, 2016). Such a perfect mapping of language knowledge onto long-term memory systems seems a perfect challenge for psychologists, and a good example of how general models of memory can be tested and advanced by the challenges posed by the hypothetical constructs present in the work of linguists.

Divjak, Milin, Medimorec, and Borowski (2022) conducted a timed grammaticality judgment experiment encompassing linguistic elements on the grammar-lexicon continuum, from syntax (subordination), via morphology (case and aspect), to lexical semantics (collocations). Starting from the established fact that concurrent working memory tasks impede what is not automatised, the authors assumed that such tasks would affect slow and conscious declarative memory but not fast and automatic non-declarative memory. Consequently, by manipulating single vs. concurrent task conditions, Divjak, Milin et al. (2022) predicted that rule-governed language items, specifically syntax and morphology, will be least affected. Their design addressed compatible models from psychology and linguistics and tested the general model of long-term memory by exploring how its subsystems are employed in accessing and retrieving language information.

The data was analysed in three different ways, considering participants' judgements in terms of Accuracy of judgment, Speed of judgment, and Consistency of speed of judgment. The experimental Condition (single vs. concurrent) did not appear to influence Accuracy, with morphological case and syntactic subordination being comparatively easier to judge (Figure 2). In terms of Speed, a significant Condition by Type of item interaction was observed, affecting linguistic Types to varying extents (Figure 3). Finally, Consistency mirrored the Accuracy pattern, revealing less variation in decision time for case and subordination items (Figure 4).

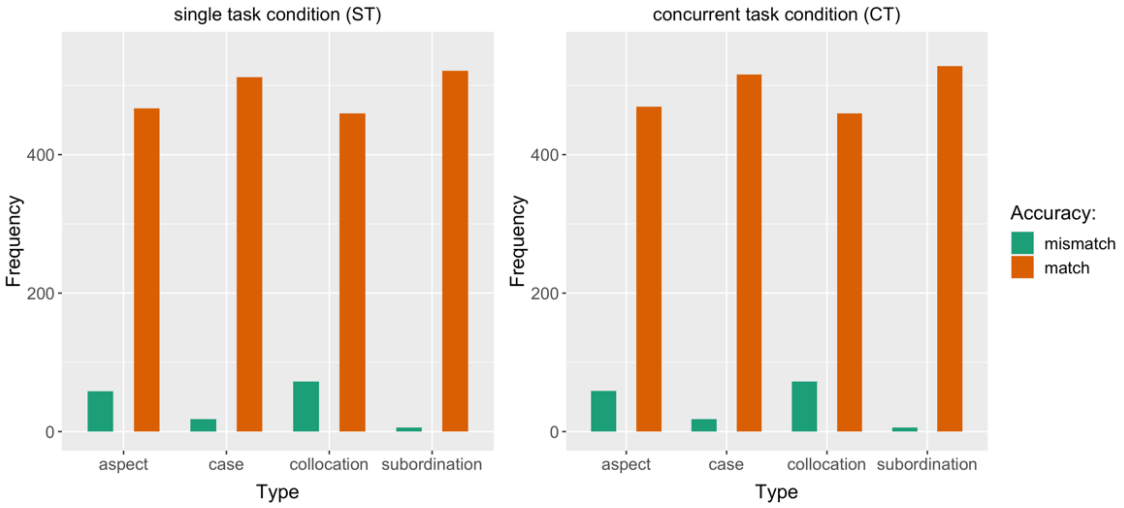


**Figure 2.** Log-linear model's predicted frequencies of Accuracy (matches and mismatches) across both experimental Conditions (adapted from Divjak, Milin et al. 2022).

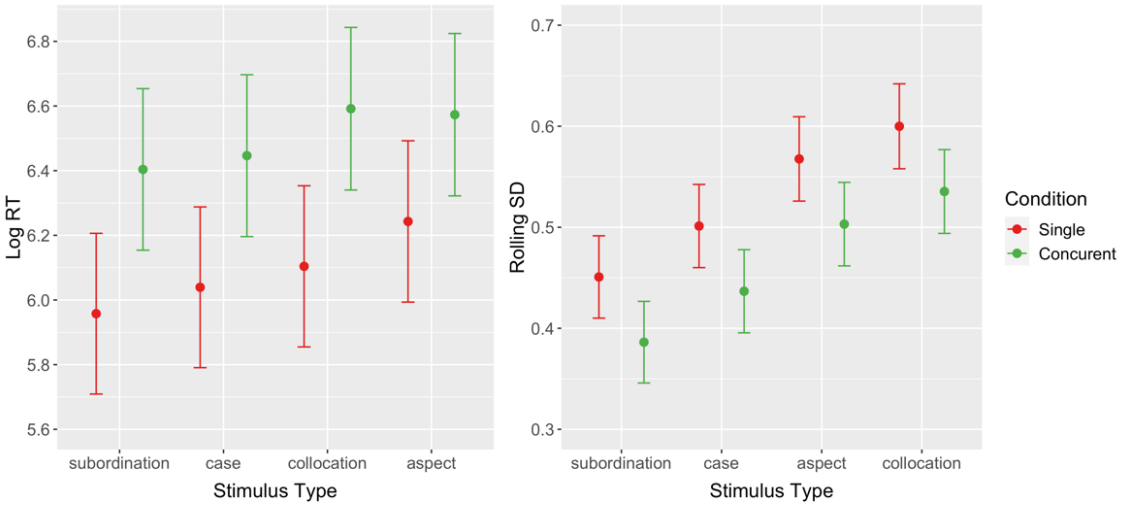


**Figure 3.** Response latencies (left) and variability in response latencies (right) for the four stimulus Types across ST and CT conditions (adapted from Divjak, Milin et al. 2022).

In summary, the experimental findings neither disprove nor confirm either of the major linguistic theories. On the one hand, they affirm the existence of the generativist's grammar-lexicon distinction, while, on the other hand, they align with the usage-based prediction suggesting a graded rather than binary nature of said distinction. Furthermore, the results contradict the generative dichotomy and support the usage-based cline with respect to the extent of proceduralisation, which seems to vary across types of linguistic knowledge. The results champion a recent, revised view of the long-term memory system, wherein the declarative and non-declarative subsystems collaborate for efficient storage and retrieval of distinct components within the same complex information package (Squire & Wixted, 2011).

**What to take home?**

Our interdisciplinary approach is centred on usage-based linguistics and the psychology of learning. Our hybrid methodology combines linguistic corpus analysis with psychological computational

modelling, providing insights into necessary and sufficient linguistic abstractions as well as necessary and sufficient learning functionality. For linguistics, this approach operationalizes the core concept of emergence through the empirically testable process of learning. Simultaneously, psychology benefits by theorizing valid outcomes for specific learning domains, and by doing so in an ecologically valid context, i.e., understanding what it takes to learn a symbolic system characterized by a complexity that far exceeds the conditions studied so far. Throughout our methodological pipeline, cognitive commitment and plausibility are rigorously pursued. We align our approach with Poggio's (2012) extension to Marr's (1982) three levels of analysis, where learning represents a self-sufficient level of explanation, influenced by the dynamic pressure of language-in-use governing what emerges and how it is learned.

**References**


Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition, 122*(3), 292-305.

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity, 2019*.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118*(3), 438-481.

Binnick, R. I. (1991). *Time and the verb: A guide to tense and aspect*: Oxford University Press.

Bouton, M. E. (2007). *Learning and behavior: A contemporary synthesis*. Sunderland: Sinauer Associates.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1959). Review of Skinner's Verbal Behavior. *Language, 35*, 26–58.

Chomsky, N. (1969). Quine's empirical assumptions. In D. Davidson & J. Hintikka (Eds.), *Words and objections: Essays on the work of W. V. Quine* (pp. 53-68). Dordrecht: Reidel.

Divjak, D., & Milin, P. (2022). *Ten Lectures on Language as Cognition: A Multi-Method Approach*. Leiden: Brill.

Divjak, D., Milin, P., Medimorec, S., & Borowski, M. (2022). Behavioral signatures of memory resources for language: looking beyond the lexicon/grammar divide. *Cognitive Science, 46*(11), e13206.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164-194.

Ellis, N. C. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition, 38*(2), 341-351.

Enquist, M., & Ghirlanda, S. (2005). *Neural networks and animal behavior* (Vol. 33): Princeton University Press.

Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.

Gardner, J. W. (1954). Foreword. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics: A Survey of Theory and Research Problems*. Baltimore: Waverly Press.

Greene, J. (1972). Psycholinguistics: Chomsky and psychology / Judith Greene. In. Harmondsworth: Harmondsworth : Penguin.

Harris, Z. (1954). Distributional structure. *Word, 10*(23), 146-162.

Haykin, S. S. (1999). *Neural networks: A comprehensive foundation*. London: Prentice Hall.

Julià, P. (1983). *Explanatory models in linguistics*. Princeton: Princeton University Press.

Leech, G. N. (1992). 100 million words of English: the British National Corpus (BNC). *Second Language Research, 28*, 1-13.

Maclay, H. (1973). Linguistics and psycholinuistics. In B. B. Kachru, H. Kahane, & R. Kahane (Eds.), *Issues in linguistics: Papers in honor of Henry and Renée Kahane*. Urbana: University of Illinois Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman & Co.

Osgood, C. E. (1968). Toward a wedding of insufficiencies. In D. TR & H. DL (Eds.), *Verbal behavior and general behavior theory* (pp. 495-519). Englewood Cliffs: Prentice-Hall.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52*(1), 111-139.

Poggio, T. (2012). The levels of understanding framework, revised. *Perception, 41*(9), 1017-1023.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science, 31*(6), 927-960.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43*(3), 151-160.

Rescorla, R. A., & Wagner, R. A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In H. Black & W. F. Proksay (Eds.), *Classical conditioning II* (pp. 64-99). New York, NY: Appleton-Century-Crofts.

Romain, L., Ez-zizi, A., Milin, P., & Divjak, D. (2022). What makes the past perfect and the future progressive? Experiential coordinates for a learnable, context-based model of tense and aspect. *Cognitive Linguistics, 33*(2), 251-289.

Roy, S., & Chakraborty, U. (2013). *Introduction to Soft computing: Neuro-Fuzzy and Genetic Algorithms*: Pearson Education India.

Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts, Inc.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory, 82*(3), 171-177.

Squire, L. R. (2009). Memory and Brain Systems: 1969–2009. *Journal of Neuroscience 29*(41), 12711-12716. doi:10.1523/JNEUROSCI.3575-09.2009

Squire, L. R., Knowlton, B. J., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44*, 453-495.

Squire, L. R., & Wixted, J. T. (2011). The Cognitive Neuroscience of Human Memory Since H.M. *Annual Review of Neuroscience, 34*, 259-288. doi:10.1146/annurev-neuro-061010-113720

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. M. R. Gabriel, J. W. Moore (Eds.), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, MIT Press, Cambridge, MA, pp. 497-537.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*(1), 231-270. doi:https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2016). The Declarative/Procedural Model: a neurobiological model of language learning, knowledge, and use. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 953-968): Academic Press.

Vygotsky, L. S. (1934/2012). *Thought and language*. Cambridge, MA: MIT press.

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. In IRE *WESCON convention record*, Vol. 4, No. 1, pp. 96-104.