

The Minimalist Program in the 2020s: Theory, Methodology, and the Road Ahead

Evelina Leivada^{1,2} & Kleanthes K. Grohmann³

*Universitat Autònoma de Barcelona*¹
*Institució Catalana de Recerca i Estudis Avançats (ICREA)*²
*University of Cyprus*³

[To appear in the *Cambridge Handbook of Minimalism*, edited by Kleanthes K. Grohmann & Evelina Leivada. Cambridge: Cambridge University Press, 2024.]

1. Introduction

The Minimalist Program (henceforth, MP or Minimalism; cf. Chomsky 1995) has been the most prominent framework in generative linguistics for over three decades. The hallmarks of a breakthrough were already evident in the early 1990s, when Chomsky & Lasnik (1993) laid out the Principles and Parameters Theory (PPT). In reality, the breakthrough had started a few years earlier, in what is often referred to as the Pisa Lectures and subsequently published as Chomsky (1981b), when the notion of ‘parameter’ was introduced in linguistic theory. Chomsky put forth the idea that an abstract grammatical system is wired in the mind of the child learner, and this innate predisposition grants a considerable aid in the process of extrapolating the target grammar.¹

The foundational pillar of MP is the idea that humans are biologically equipped with the capacity to acquire any natural, spoken or signed, language because Language is innate to humans. A key notion within MP is the initial state of this predisposition for Language, called Universal Grammar (UG) by Port-Royal grammarians. Succinctly put, UG “may be thought of as some system of principles, common to the species and available to each individual prior to experience” (Chomsky 1981a: 7). Establishing a link between UG and language acquisition, Chomsky further defined UG, in a typical acquisition scenario, as “a characterization of children’s pre-linguistic initial state” (p. 7).

¹ The fundamental insight of PPT—starting with Government and Binding Theory (Chomsky 1981b, 1986)—underlying the minimalist enterprise forms the centerpiece of the MP textbook introduction by Hornstein et al. (2005). More historical roots are detailed especially in Part I of the present handbook and various chapters in Part I of Leivada & Grohmann (2024).

The specifics of these claims—innateness, the degree of language specificity, and the role of general cognition—have been a matter of long and intense debates (for recent overviews, see e.g. Roberts 2016, MENDÍVIL-GIRÓ 2018, Grohmann & Leivada 2020). Given the magnitude of Chomsky’s influence in science across different fields of research, it is perhaps unsurprising that the debates continue as lively today as they were a couple of decades ago. A good 30 years into Minimalism (Chomsky 1993), its theses have been so intensely scrutinized over and over, that neither novel support nor untried criticism are possible at this stage. What can be examined and (re)synthesized are the theoretical and empirical offsprings of MP, understood as potentially convergent or partly divergent lines of research that share roots which go back to Chomsky’s work.

Employing this perspective, we aim to provide an update on the research questions that guide generative linguistics as well as on the methodologies that synthesize the evidential basis of the field, reflecting three crucial dimensions: (i) the importance of having a theory, (ii) the relation between theory and measurements, and (iii) the results this coupling of theory and measurements has produced in the last three decades of generative research. Putting these three dimensions together, our discussion examines the answers MP and the research that was developed on its basis offer with respect to how Language works (behavior), through the experimental lens we have on its operation (observables), in the form of a principled approach to what has been measured, why so, and what are the adduced explanations (theory).

2. What is a theory and does MP have one?

Theories are bodies of knowledge that are broad in scope involve interrelated concepts, definitions, and propositions with the aim to explain phenomena and offer predictions (Glanz 2008). Perhaps the most frequent strand of criticism towards Minimalism concerns its status as a program and its alleged absence of theories. This criticism comes in many guises, but most often it assumes the form of concerns about (not) making predictions, (not) offering falsifiable hypotheses, and (mis)presenting explananda as part of theory-internal explanations (Tomasello 1995, Dąbrowska 2015).

As its name suggests, MP is a *program*, not a theory (Chomsky 1995, Freidin 1997). How long does it take for a program to develop and test a full-fledged theory? This question is missing something important: MP already has many theories, some in early stages of their development, others more thoroughly explored.² To name just a few, UG makes provisions for the workings of various subsystems including X-bar Theory, Binding Theory, Case Theory, Theta Theory, Bounding Theory, and PPT itself, the theory of principles and parameters (Chomsky 1986). Importantly, minimalist landmark theories make falsifiable predictions and offer experimentally tested or testable hypotheses. For instance, many of these theories build on the claim that human

² Besides, as Chomsky has repeatedly said, there is nothing ‘special’ about Minimalism, neither the research program nor the term itself. Consider this quote (Chomsky 2011: 10):

In recent years, the inquiry into these topics is often called the minimalist program, but that term has been extremely misleading. I’m sorry I ever used it. It’s just ordinary science, normal ordinary science, extending the main thrust of theoretical linguistics since the earliest days of contemporary work in the nineteen-fifties.

language can encode both local and non-local abstract grammatical dependencies. Recent experiments confirm that this claim is correct, and this ability is already developed in eighteen-month-old infants (Perkins & Lidz 2021). Similarly, the existence of prototypical thematic roles such as Agent and Patient, put forth within Theta Theory, has been experimentally supported by many studies of event cognition and sentence comprehension (Rissman & Majid 2019). In Bounding Theory, varying constraints on movement have been confirmed in all tested languages and several grammar-, discourse-, or resource-based explanations have been adduced for them (Boeckx 2012, Phillips 2013).

Last, PPT, perhaps the most criticized theory of all the theories proposed within MP, does not run into the recurrently brought up problem of unfalsifiability (cf. Evans & Levinson's 2009 claim of UG being immune to falsification): Computational and other types of experimental analyses of UG parameters have found both strengths and challenges in PPT, further proposing rectifying amendments to the theory (Boeckx & Leivada 2013, Leivada et al. 2017). For example, one important modification boils down to variation progressing from being linked to a range of innate, syntactic parameters in the early stages of the theory (Chomsky 1981b) to nowadays being conceived as neither syntactic (cf. Berwick & Chomsky 2011, who link variation to the externalization components of language) nor innate (at least not in the form of UG encoded parameters; Biberauer et al. 2014). Other theories explored within the generative framework were found to be empirically robust, hence they were not significantly amended (such as the critical period of language acquisition or the distinction between a model for competence and a model for performance).

The take-home message is that if falsifiability is defined as the ability to subject theories as well as their predictions to analyses that confirm or disconfirm them, proposals about the primitives of UG have already been put to test, hence there exist ways of falsifying these proposals. From this perspective, asking “what kind of evidence could refute the UG *hypothesis*” (Tomasello 2007, emphasis added) simply misstates the facts. UG is neither a hypothesis nor a theory—it is a concept within a theory (of innateness), developed within a program (MP).

Having established that MP is not a theory but a program of study that involves many falsifiable theories, it is important to understand the role of theory in the field as well as the caliber of the various existent theories in this specific framework. If one attempted to draw a circle of theory-formation in natural sciences, they would probably arrive at a schematic representation like the one given in figure 1.

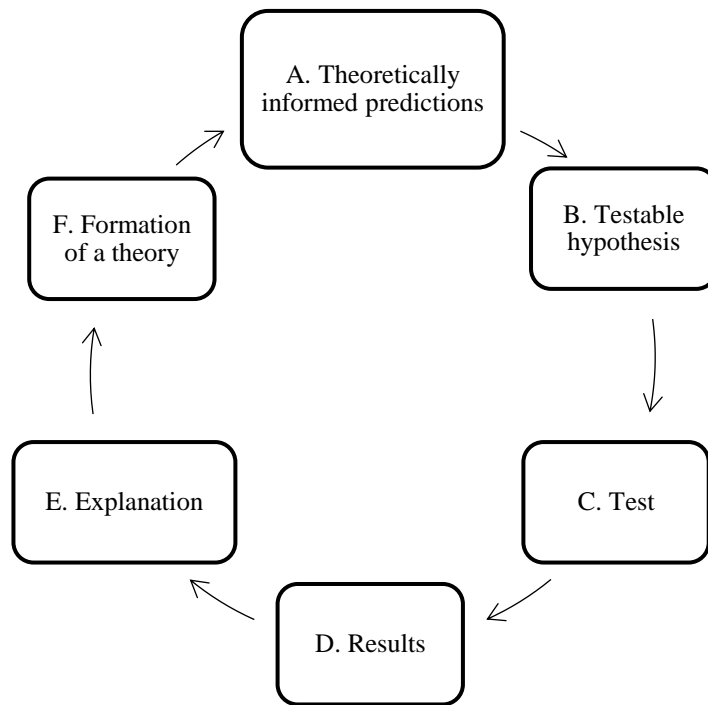


Figure 1: The process of theory-formation

For example, in biology journals that are devoted to theoretical matters, the published works often advance predictions that are formulated in a way that is amenable to empirical (dis)confirmation. Depending on whether the obtained results confirm them or not, the best sets of explanations are then synthesized into theories. In other words, the starting point is not the theory (step F in figure 1), but a set of theoretically informed hypotheses (step A) that crystallize into more narrowly presented research questions and/or hypotheses (step B), which are then put to test (step C). Even though fragments of a theory are present in step A, the formation of the relevant theory comes last and is a necessary step in the process.

In fact, robust results and their explanations may be questioned precisely due to the fact that at times they are not embedded in and explained under an overall theory. To give an example, today we know that the results of Semmelweis on the role of germs in the incidence of puerperal fever were correct. However, at the time of their discovery, they were rejected by the established medical community in part because they lacked an explanation that went significantly *beyond what was observed*: They did not come as part of a theory that explained what happened and why. Semmelweis' results were confirmed, enjoying widespread acceptance, some decades later when Pasteur formulated the germ *theory* of disease. Undoubtedly, any set of results can be evaluated quite differently in the presence or absence of an overall theory that connects the various dots.

Bringing linguistics into this picture, a unique selling point of MP (often unacknowledged even by linguists) is precisely its richness in theory. Natural sciences are often highly focused on experiments, conditions, and measurements, while at times lacking a principled approach to what is being measured, why so, and what this means in relation to a specific theory. Most work in

social sciences looks like an exercise in advanced statistical model fitting, which often falls short of theory building too (Fried 2021). Similarly, certain branches of psychology show an increased interest in measurements, co-efficients, and other statistical paraphernalia, without always placing an equal weight on theory. As Cummins (2000: 120) observes: “In psychology, we are overwhelmed with things to explain, and somewhat underwhelmed by things to explain them with.”

Generative linguistics differs and has differed in this respect since its very conception. In *Aspects of the Theory of Syntax*, Chomsky (1965) highlighted that the key issues which needed to be addressed first concerned the development of a theory for the faculty of language, even if this theory relied on certain idealizations (e.g., the notion of an ideal speaker–listener who knows their language perfectly) that removed actual measurements and other experimentally-driven concerns from the picture.³ In other words, *Aspects* presented theoretically informed, tentative assumptions, and in doing so Chomsky put to practice the dictum Richard Feynman formalized in his lectures about the scientific method around that time: You start with theoretically informed guesses (Feynman 1964). This view is in line with the claim that there are no atheoretical approaches to linguistic data (Ludlow 2019). Traces of some theory are necessarily present in step A and connections of the new/revised theory presented in step F to existing theories should exist.

Arguing that research in MP has always been theory-driven does not entail that all the theories that have ever been entertained within this program are by default (equally) strong. The claim we put forth is that MP *is* theory-rich and most of its theories *are* falsifiable. As happens in all fields of science—and in generative linguistics, too—some theories are stronger than others, and some theories have been confirmed through robust experimental evidence more than others. Employing Fried’s (2021) definitions, we make a distinction between strong and weak theories. The former present an unambiguous set of assumptions about a phenomenon as well as an explanation of it that goes beyond a jargon-loaded restatement of the explananda/observable data. Strong theories make predictions (e.g., Einstein’s theory of general relativity predicted the existence of gravitational waves, the existence of which was confirmed a century later; Abbott et al. 2016) and can be applied in hypothetical scenarios (e.g., we have theories about what happens to skyscrapers during earthquakes that enable us to estimate the outcomes of a hypothetical earthquake in relation to a specific skyscraper, in the absence of an actual earthquake; Fried 2021).

³ More specifically, Chomsky argued:

All such theories therefore require supplementation by an evaluation measure if language acquisition is to be accounted for and selection of specific grammars is to be justified; and I shall continue to assume tentatively, as heretofore, that this is an empirical fact about the innate human *faculté de langage* and consequently about general linguistic theory as well. (Chomsky 1965: 37)

In addition:

In linguistics, it seems to me that sharpening of the data by more objective tests is a matter of small importance for the problems at hand. [...] Perhaps the day will come when the kinds of data that we now can obtain in abundance will be insufficient to resolve deeper questions concerning the structure of language. (Chomsky 1965: 20-21).

Weak theories neither offer real explanations nor make such predictions. Also, they do not spell out the conditions under which a phenomenon is expected to be found.

It has been observed that theories in (clinical) psychology are often of the weak type (Fried 2021). If one agrees with Chomsky's (2007) assessment that there is no other way to conceive of linguistics but as being a part of psychology, one may wonder what type of theories are found in this subpart of the discipline. We argue that, unsurprisingly, MP involves both types. On the one hand, there are strong theories, which have been articulated in detail, allowed for the formation of falsifiable predictions, and, when necessary, have been reformed (e.g., the earlier mentioned example of principles and parameters). Importantly, when a theory allows for reform, it does not mean that it is a weak theory. On the other hand, some MP theories are weak for they lack an explanation altogether. Presenting the data and describing a phenomenon—say, sentence X in language Y instantiates a phenomenon Z, and according to the observable data, Z has the following properties A, B, C—in a way that does not involve explanations or predictions does not constitute a strong theory.

Three important points are necessary in order to clarify the scope of this point. First, this claim is not made specifically in relation to MP, PPT, or even generative linguistics at large. Weak theories are pervasive across linguistic frameworks and across scientific fields in general. Second, in agreement with our earlier discussion of step A in figure 1, we do not argue that presentations and descriptions of data in MP are theory-free (cf. Haspelmath 2010, 2021 and Ludlow 2019 on diverging views on this point). We argue that mere presentations of data do not constitute strong theories, even though they may provide the evidential basis on which strong theories may develop. A grammar of language L is a theory of L, as Chomsky (1957: 49) pointed out—but it is not a strong theory unless it involves predictions and explanations. In *Syntactic Structures*, Chomsky (1957) highlights the importance of *explanatory power* of a theory, when he argues that it is the linguist's task to describe linguistic phenomena in terms of different levels of representation, “some of which are quite abstract and non-trivial” (p. 85). Third, the overall status of (weak) theories as theories is debatable and depends on the definitions one endorses.⁴

If one defines ‘theory’ as a set of interrelated observations for a phenomenon—a definition very different from the one that opened this section—any description of a grammatical phenomenon X in language L is a theory of X, regardless of whether it is limited on observations of tangible data or not. This, however, is not the definition of the theory that most scientific fields endorse. Not drawing the distinction between strong and weak theories poses the risk of mistaking explananda (i.e. phenomena that have been observed and are in need of explanations and theoretical synthesis) for explanations, consequently presuming that a discipline fares well on the explanatory front when it may not. To repeat, this is not a problem we attribute to theoretical linguistics or MP specifically. In fact, it is entirely conceivable that theoretical linguistics may fare better than other branches of psychology in this respect: By not being so strongly preoccupied with statistical modelling and by permitting certain types of idealizations that sidestep experimental

⁴ To understand the meaning of the parenthetical remark, see Fried's (2021) view that weak theories should not be called theories in the first place, as they may impede theory formation, failure, and reform.

concerns, it reserves more space for theory—certainly so in terms of intentions, but possibly in terms of actual practice too.

3. Methodology in theoretical linguistics

Judgments of acceptability largely form the evidential basis of linguistics. A typical criticism addressed towards generative grammar (although it should be more fittingly addressed towards all frameworks that use introspective judgments) has to do with the methodologies employed to collect the data that support its theories. For example, Featherston (2007: 278) argues:

It is simply inadequate research practice for linguists to rely on their own unconfirmed introspective judgements as linguistic evidence. It is also insufficient to take judgements uncritically from the literature and build upon them, since the literature is full of examples of dubious practice. I have on my desk a paper by an author with a PhD from MIT and a job at a prestigious university. The main point of the article rests upon the interpretation of two structures (both in English and German). But the author is not a native speaker of either of these languages and these critical readings are simply unavailable to me or anyone I have asked. Some related structures do show the intended readings, as do the original examples given in the first paper on the issue 25 years ago, but the author has simply not checked the facts. This is not a new problem. Greenbaum (1977) comments “All too often the data in linguistics books and articles are dubious, which in turn casts doubt on the analyses. Since analyses usually build on the results of previous analyses, the effect of one set of dubious data left unquestioned can have far-reaching repercussions that are not easily perceived.”

The magnitude of the alleged discrepancy between informally elicited judgments and judgments elicited through proper experiments has been recently quantified. The reliability of informally elicited acceptability data has been repeatedly demonstrated, at least when it comes to official languages with millions of speakers like English (Sprouse & Almeida 2012, Sprouse et al. 2013). Experimental replications of acceptability judgments from other languages did not show equally high percentages of agreement between judgments obtained in informal ways vs. controlled experiments (Linzen & Oseki 2018), but such differences probably have to do with the individual characteristics of each language in terms of variation. While various reasons can explain this diminished degree of agreement in linguistic communities that involve small and/or non-official languages (Leivada et al. 2019), it is important to address some common misconceptions about the evidential basis of MP.

The first one has to do with the quantity of the data that are elicited informally. Let’s take, for instance, the claim that “[*m*]ost acceptability judgments reported in the syntactic literature are obtained by linguists being their own informants” (Juzek & Häussler 2020: 109, italics in original). With respect to the syntactic literature that discusses big languages, perhaps this claim provides an accurate description of how the landscape of theoretical linguistics looked some decades ago. Today, however, most published work in generative linguistics either puts any informally elicited

data in comparison to data presented in previous studies—hence dubious data are very likely to be eventually called out—or presents quantitative data, involving methodologies and techniques that go beyond informal judgments.

The second misconception boils down to the assumption that all linguistic communities are equally uniform in terms of variation. Although we cannot confirm whether most of the recently reported data are elicited through informal methods or not, if we assume—for the sake of the discussion—that they are, the repercussions of unsystematic data collection are real in the context of small and/or non-official languages, where the linguistic landscape may hide a greater degree of variation (Leivada et al. 2019). This is a delicate methodological issue. On the one hand, presenting data from a single speaker is useful and possibly even necessary, when the linguistic community poses limits (e.g., in the case of moribund varieties; D’Alessandro et al. 2021). Succinctly put, we should not contribute to the marginalization of small linguistic communities—which may not permit conducting large-scale studies—by not engaging in small-scale, exploratory research that helps promote their visibility and representation in the overall scientific arena. On the other hand, if the tested community is very heterogeneous, one should ensure that the data obtained informally from an $n=1$ sample is representative of the language of the community.

This is a frequently overlooked point with grievous consequences for the evidential basis that supports the theory. Usually, linguists do not present their findings as making claims about specific individuals, but about languages and/or linguistic communities. In other words, data from $n=1$ is never framed as making claims about one speaker/signer; it is construed more broadly in the form of claims about a language. This leap may or may not be justified. In the second scenario, if the judgments from this one speaker/signer are not representative of the language of their community, and the linguist presents them as if they were, a false effect finds its way to the published record. Controlled experiments may also report false effects (and in fact they often do; Ioannidis 2005), but at least they explain the conditions under which an effect was found, hence they permit replication. If replication repeatedly fails, the effect will be flagged as false or dubious. If, however, the road to the effect is not described, it is not that easy to determine its *ecological validity*.

One response to this concern is that obtaining observations from one speaker/signer is still informative as to what the faculty of language permits, hence it can act as a reliable and valuable proxy to the latter (D’Alessandro et al. 2021). This observation is correct, but it comes with two caveats. First, this observation does not change the fact that if the data behind a described effect is representative only of one speaker/signer, and instead of being flagged as such, it is presented in the form of claims about a language, a false effect *is* reported. Second, even if the informally gathered data from $n=1$ is representative of the community, one should ponder what kind of evidence this is. Asking a friend in an informal context whether sentence X sounds fine is not equally reliable as tapping into the acceptability of X through a well-powered, controlled experimental study. In the field of health sciences, Belluz & Hoffman (2015) discuss the relation that exists between the type of study and the strength of conclusions that the study permits. Case studies, which are described as the weakest kind, are found at the bottom of the hierarchy, while

meta-analyses and randomized controlled experiments occupy the highest positions in the hierarchy of strength of evidence. Similarly to how theories are not all equal in terms of strength, studies are not all equally reliable, even when they adduce evidence for the same effect. Going back to linguistics, already in *Aspects*, Chomsky foresaw the possibility that someday informally elicited judgments would not be sufficient to resolve the deeper questions concerning the structure of language (Chomsky 1965: 21). As Schütze (2016) shows, the day has come.

From a methodological point of view, should we then not test small and/or non-standard languages using informal elicitation techniques? D'Alessandro et al. (2021) convincingly explain that we should. At the same time, we should flag informally elicited data with unverified ecological validity as such. Exploratory research should be explicitly presented as such, too, and unwarranted generalizations should be framed as working hypotheses. Theories can only be as good as the data that support them. Both MP and its latest precursors have been mindful of this and of the overall importance of *reliable* empirical data in defining and refining the principles of UG. This is, as Chomsky (1967) argued in his appendix to Eric Lenneberg's *Biological Foundations of Language*, the central and critical problem for linguistics.

4. The road ahead

The Cambridge Handbook of Minimalism aims to provide a comprehensive guide to present developments in generative syntactic theory typically placed under the umbrella 'Minimalist Program'. It is structured to include (i) the historical context and foundation of the program, providing the big picture, (ii) a survey of the variety of phenomena dealt within Minimalism in terms of concepts and relations, and (iii) a part dedicated to operations and processes. This handbook is thematically complemented by a second compendium, *The Cambridge Handbook of Minimalism and Its Interdisciplinary Applications* (Leivada & Grohmann 2024), which provides an overview of the major areas of research within modern syntactic theory, and an engagement of the interfaces between MP and other domains of linguistic inquiry and cognitive science. The latter involves contributions that show how core concepts and outputs of Minimalism can be approached in a way that goes beyond the narrow research aims of theoretical linguistics. The two handbooks jointly cover a wide variety of (sub)disciplines, ranging from language development and psycho-, neuro-, and clinical linguistics to biology, evolution, animal cognition, computational linguistics, and artificial intelligence.

In terms of results, we suggest that these contributions demonstrate that some MP concepts are of the right granularity to enable the development of linking hypotheses across different subfields of study in psychology and neuroscience. These linking hypotheses are precisely what is necessary in order to address the *mapping problem*, which refers to the relation between the primitives of cognition (in this case, language) and neurobiology (Poeppel 2012). Overall, such discussions pave the way for discussing whether a negative answer to the *ontological incommensurability problem* (according to which the units of linguistic computation and the units of neurological computation are incommensurable; Poeppel & Embick 2005) is possible. It is likely that the road ahead entails pursuing comparative studies of the 'parts list' of different

disciplines, something that can help restore the influence and central position of linguistics in the general scheme of things in the field of cognitive science (Grohmann & Leivada 2020).

Bringing other areas of psychology, biology, computation, and evolution into the picture, while at the same time preserving our focus on linguistic concepts and research questions, is probably the only way of achieving the true purposes of Minimalism, as outlined in Chomsky (1995): to go beyond intuitive accounts that seem simple and persuasive, but may be inadequate, and to recognize the unsuspected richness of the phenomena of language that this framework investigates.

References

- Abbott, B. P. et al. 2016. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters* 116: 061102.
- Belluz, J. & S. Hoffman. 2015. The one chart you need to understand any health study. <https://www.vox.com/2015/1/5/7482871/types-of-study-design>
- Berwick, R. C. & N. Chomsky. 2011. The biolinguistic program: The current state of its development. In A. M. Di Sciullo & C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 19-41. Oxford: Oxford University Press.
- Biberauer, T., I. Roberts & M. Sheehan. 2014. No-choice parameters and the limits of syntactic variation. In R. Santana-LaBarge (ed.), *WCCFL 31: Proceedings of the 31st West Coast Conference on Formal Linguistics*, 46-55. Somerville, MA: Cascadilla Press.
- Boeckx, C. 2012. *Syntactic Islands*. Cambridge: Cambridge University Press.
- Boeckx, C. & E. Leivada. 2013. Entangled parametric hierarchies: Problems for an overspecified Universal Grammar. *PLoS ONE* 8(9): e72357.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1967. The formal nature of language. In E. Lenneberg, *Biological Foundations of Language*, 397-442. New York: John Wiley & Sons.
- Chomsky, N. 1981a. On the representation of form and function. *The Linguistic Review* 1, 3-40.
- Chomsky, N. 1981b. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
- Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York, NY: Praeger.
- Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1-52. Cambridge, MA: MIT Press. [Reprinted in Chomsky (1995), 167-217.]
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2007. Approaching UG from below. In U. Sauerland & H.-M. Gärtner (eds.), *Interfaces + Recursion = Language?*, 1-29. Berlin: Mouton de Gruyter.
- Chomsky, N. 2011. Language and the cognitive science revolution(s). Transcribed text of lecture given at Carleton University, 8 April 2011.

- Chomsky, N. & H. Lasnik. 1993. The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann (eds), *Syntax: Ein internationales Handbuch zeitgenössischer Forschung — An International Handbook of Contemporary Research*, vol. I, 506-570. Berlin: Mouton de Gruyter. [Reprinted in Chomsky (1995), 13–127.].
- Cummins, R. 2000. “How does it work?” versus “What are the laws?”: Two conceptions of psychological explanation. In F. C. Keil & R. A. Wilson (eds.), *Explanation and Cognition*, 117-144. Cambridge, MA: MIT Press.
- Dąbrowska, E. 2015. What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology* 6: 852.
- D’Alessandro, R., D. Natvig & M. T. Putnam. 2021. Addressing challenges in formal research on moribund heritage languages: A path forward. *Frontiers in Psychology* 12: 700126.
- Evans, N. & S. C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32, 429-448.
- Featherston, S. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33(3), 269-318.
- Feynman, R. 1964. Lectures on the scientific method. Available at <https://www.youtube.com/watch?v=OL6-x0modwY>
- Freidin, R. 1997. *The Minimalist Program* by Noam Chomsky. *Language* 73(3), 571-582.
- Fried, E. I. 2021. Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory* 31(4), 271-288.
- Glanz, K. 2008. *Health Behavior and Health Education: Theory, Research, and Practice*. San Francisco, CA: Jossey-Bass Inc.
- Greenbaum, S. 1977. The linguist as experimenter. In Fred R. Eckman (ed.), *Current Themes in Linguistics: Bilingualism, Experimental Linguistics, and Language Typologies*, 125-144. Washington, DC: Hemisphere.
- Grohmann, K. K. & E. Leivada. 2020. Reconciling linguistic theories on comparative variation with an evolutionarily plausible language faculty. In A. Bárány, T. Biberauer, J. Douglas & S. Vikner (eds.), *Syntactic Architecture and Its Consequences II: Between Syntax and Morphology*, 25-42. Berlin: Language Science Press.
- Haspelmath, M. 2010. Framework-free grammatical theory. In B. Heine & H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 341-365. Oxford: Oxford University Press.
- Haspelmath, M. 2021. General linguistics must be based on universals (or non-conventional aspects of language). *Theoretical Linguistics* 47(1-2), 1-31.
- Hornstein, N., J. Nunes & K. K. Grohmann. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press.
- Ioannidis, J. P. A. 2005. Why most published research findings are false. *PLoS Medicine* 2: e12.
- Juzek, T. S. & J. Häussler. 2020. Data convergence in syntactic theory and the role of sentence pairs. *Zeitschrift für Sprachwissenschaft* 39(2), 109-147.

- Leivada, E., E. Papadopoulou & N. Pavlou. 2017. Functionally equivalent variants in a non-standard variety and their implications for Universal Grammar: A spontaneous speech corpus. *Frontiers in Psychology* 8: 1260.
- Leivada, E., R. D'Alessandro & K. K. Grohmann. 2019. Eliciting big data from small, young, or non-standard languages: 10 experimental challenges. *Frontiers in Psychology* 10: 313.
- Leivada, E. & K. K. Grohmann (eds.). 2024. *The Cambridge Handbook of Minimalism and Its Interdisciplinary Applications*. Cambridge: Cambridge University Press.
- Linzen, T. & Y. Oseki. 2018. The reliability of acceptability judgments across languages. *Glossa: A Journal of General Linguistics* 3(1): 100.
- Ludlow, P. 2019. The philosophy of generative linguistics: best theory criteria. In A. Kertész, E. Moravcsik & C. Rákosi (eds.), *Current Approaches to Syntax*, 521-548. Berlin: Mouton de Gruyter.
- Mendívil-Giró, J. L. 2018. Is Universal Grammar ready for retirement? A short review of a longstanding misinterpretation. *Journal of Linguistics* 54(4), 859-888.
- Perkins, L. & J. Lidz. 2021. Eighteen-month-old infants represent nonlocal syntactic dependencies. *PNAS* 118(41): e2026469118.
- Phillips, C. 2013. On the nature of island constraints I: Language processing and reductionist accounts. In J. Sprouse & N. Hornstein (eds.), *Experimental Syntax and Island Effects*, 64-108. Cambridge: Cambridge University Press.
- Poeppel, D. 2012. The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology* 29(1-2), 34-55.
- Poeppel, D. & D. Embick. 2005. Defining the relation between linguistics and neuroscience. In A. Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103-118. Mahwah, NJ: Lawrence Erlbaum.
- Rissman, L. & A. Majid. 2019. Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review* 26, 1850-1869.
- Roberts, I. Introduction. In I. Roberts (ed.), *The Oxford Handbook of Universal Grammar*, 1-34. Oxford: Oxford University Press.
- Schütze, C. T. 2016. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Berlin: Language Science Press.
- Sprouse, J. & D. Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*. *Journal of Linguistics* 48, 609-652.
- Sprouse, J., C. Schütze & D. Almeida. 2013. Assessing the reliability of journal data in syntax: *Linguistic Inquiry* 2001–2010. *Lingua* 134, 219-248.
- Tomasello, M. 1995. Language is not an instinct. *Cognitive Development* 10, 131-156.
- Tomasello, M. 2007. What kind of evidence could refute the UG hypothesis? Commentary on Wunderlich. In Martina Penke & Anette Rosenbach (eds.), *What Counts as Evidence in Linguistics*, 175-178. Amsterdam: John Benjamins.