Spring 2019

Tool Migration: A Framework to Study the Cross-disciplinary Use of Mathematical Constructs in Science

Chia-Hua Lin

Recommended Citation

Lin, C.(2019). *Tool Migration: A Framework to Study the Cross-disciplinary Use of Mathematical Constructs in Science.* (Doctoral dissertation). Retrieved from https://scholarcommons.sc.edu/etd/5170

Tool Migration: A Framework to Study the Cross-disciplinary Use of Mathematical Constructs in Science

by

Chia-Hua Lin

Bachelor of Arts
Soochow University 2008

Master of Arts
National Yang-Ming University 2010

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Philosophy

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

Michael Dickson, Major Professor

Colin Allen, Committee Member

Anne Bezuidenhout, Committee Member

Tarja Knuuttila, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

Dedication

To all the brave scientists.

Acknowledgments

I would like to express my deepest appreciation to my dissertation committee: my advisor Michael Dickson, who tirelessly supported me during this process and through the evolution of my scholarship; Tarja Knuuttila, who counseled me each time I came to a crossroad; Colin Allen, who encouraged me to keep moving forward, and Anne Bezuidenhout, who challenged me to become a better philosopher. I am indebted to the Office of the Vice President for Research at the University of South Carolina and the Konrad Lorenz Institute for providing financial support for this project. I would also like to thank Paul Humphreys, Guido Ganiglia, Lynn Chiu, Ivan Gonzalez Cabrera, Luis Alejandro Villanueva Hernandez, Thomas Reydon, Mary Morgan, Julia Bursten, and Thomas Pradeu for providing critically insightful feedback on my work. Particularly helpful to me during this time was my Family of Choice: Mindi Spencer, John Raines, Erin Roberts, and my four-legged friends - Loki, Snowbelle, and Penny. Together, they saw to both the completion of my work and the upkeep of my well-being. As someone who stumbled into philosophy as an undergrad, I am grateful for Chienkuo Mi, Wen-fang Wang, and Allen Houng for believing that I could not only take, but successfully complete this great adventure. Most importantly, I cannot be more appreciative to Seo Yeon Park, who always inspires me to be the best version of myself.

Abstract

This dissertation is concerned with the scientific practice in which a mathematical construct that was originally developed to study a particular subject matter subsequently used in other disciplines or sub-disciplines for a different subject matter, a phenomenon that I call 'tool migration.' I argue that tool migration can be 'epistemically risky.' Specifically, uprooting a research tool from one disciplinary context and re-situating it for use in another can change how the tool is applied; whatever has made the tool useful and reliable in the first place may not have stayed the same in the new context. Using the migrations of game theory and formal language theory as examples, I identify three kinds of epistemic risks associated with tool migration: mischaracterization (i.e., characterizing a new phenomenon with inappropriate assumptions), misinterpretation, (i.e., using ill-fitted background contexts to interpret the result), and misjudgment (i.e., incorrectly rejecting or accepting a novel use of a migrating research tool). However, my analyses of these tool migration stories show that (1) proactive modifications to a tool in migration are conducive to successfully applying an old tool in a new context, (2) being aware of the changes to a tool due to migration is crucial to avoid misinterpretation and misjudgment, and (3) there is a need for a study of tool migration to understand how scientists manage (or may manage) these risks.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AGL	Artificial Grammar Learning
ALGOL	Algorithmic Language
BNF	Backus-Naur Form
CGT	Classical Game Theory
EGT	Evolutionary Game Theory
ESS	Evolutionarily Stable Strategy
FLB	The Faculty of Language in the Broad Sense
FLN	The Faculty of Language in the Narrow Sense
FLT	Formal Language Theory
IC	Immediate Constituent
IFG	Inferior Frontal Gyrus
NP	Noun Phrase
PSG	Phrase Structure Grammar
VP	Verb Phrase

Chapter 1 Introduction

1.1 Orientation

Research projects and programs involving researchers from multiple disciplines are a growing trend as part of scientific practice. As a result, inter-, cross-, or multi-disciplinary research programs have become a significant force in knowledge generation. One important aspect of this trend in knowledge production is the development and use of formal models, theories, or methods that are meant for bridging gaps between scientists who have different training and approaches.

An example of such an emerging research program is *cognitive biology*, one strand of experimental psychology that emerges from the development of the theory of formal languages (or formal language theory 'FLT') in linguistics and computer science.[1] FLT, including its components such as automata theory and the Chomsky hierarchy, is the study of mathematically defined languages. Cognitive biology is the study of cognition as a biological function. One advocate of this new strand of cognitive science, Tecumseh Fitch (2014, 330), argues that a "triangulation between three disciplines" neuroscience, the cognitive sciences, and cognitive biology is required to properly understand both the evolution of and neural substrate for human linguistic capacity. According to Fitch (ibid., 330), neuroscience provides an understanding of "the firm physical foundations of brain

---

[1] Throughout the dissertation, I will use 'the theory of formal languages' and 'formal language theory' interchangeably.

function," cognitive biology provides "a comparative viewpoint." Most germane to my inquiry, Fitch proposes that the "bridging functions between all" these "domains should build upon the insights of computer science" and, in particular, on the theories of computability and FLT (ibid., 330).

Fitch's proposal would be familiar to the philosophers of mind, as it is rooted in the classical, computational theory of mind, which views the human mind or the human brain (or, to some philosophers, both) as an information processing system. In the traditionally anthropocentric research landscape of cognitive science, what is new in Fitch's approach to cognition is the stress on the inclusion of nonhuman animal model organism in the study of human cognition.[2] For Fitch, essential to the inclusion is the cross-disciplinary application of the theory of formal languages. Applying FLT allows experimental psychologists to design experiments to probe the abilities of subjects—both human and nonhuman animals—to learn artificial grammars. The experimental results, in turn, shed light on 1) which kind of information processing system resides in the subjects and 2) the differences between the species of subjects in terms of the information processing system that they possess.

In this dissertation, I focus primarily on the journey of the theory of formal languages. FLT was originated in Chomsky's study of natural languages (Chomsky 1956; Ginsburg 1980). Between its initial formulation by Chomsky and its recent, novel applications in cognitive biology (e.g., Fitch and Hauser 2004), FLT received intervening development in computer science for improving the design of both the programming

---

[2] Early exceptions include, e.g., Brooks 1991; Bechtel, Graham, and Balota 1998, chap. 8; Bekoff, Allen, and Burghardt 2002; Allen and Bekoff 1999. Some of the more recent work includes, e.g., Andrews 2014.

languages and their corresponding recognizers, called ‘automata’ (e.g., Greibach 1981). Since 2000, the results of artificial grammar learning experiment have been used to infer about the computational constraints to the subjects' cognitive infrastructures (e.g., Hauser, Chomsky, and Fitch 2002; Fitch and Hauser 2004; ten Cate and Okanoya 2012). The use, including the experimental use, of FLT in cognitive biology thus presents a rich source for understanding knowledge production in the era of inter- and cross-disciplinarity — the subject matter that I set out to explore in this dissertation.

1.2 Motivation

The story of FLT makes a curious case to philosophers of science, particularly those to whom cross-disciplinarity as scientific practice is of interest. How has a theory developed in linguistics for studying natural languages become the backbone of computing technologies and now a promising bridge, as Fitch argues, that links cognitive biology, neuroscience, and the cognitive sciences?

Nonetheless, outside its native and home disciplines (linguistics and computer science, respectively), FLT, just like many other theoretical frameworks and models, has been applied in various kinds of research. In addition to cognitive biology that I am investigating in this dissertation, the theory has been applied by molecular biologists to tackle various problems, including the problem of recognizing protein-encoding genes (e.g., Head 1987; Dong and Searls 1994; Searls 1992, 1993, 1995, 2002), the implementation of computation in cells or DNA (e.g., Paun et al. 1998/2005), and the design and verification of synthetic genetic constructs (e.g., Cai et al. 2007; Coll et al. 2016).

In light of these various areas that the theory participates, I term my approach to the case of FLT a study of ‘tool migration’ - a project that investigates how formal constructs such as theoretical frameworks or mathematical models ‘migrate’ across disciplinary contexts to gain new ways of using them. If one views scientific research as knowledge production of some sort, then mathematical constructs like FLT can be seen as playing crucial, facilitating roles in the overall generation of knowledge. Moreover, if one thinks of these facilitating theories or models in science as ‘tools for research,’ which is a position I will defend in Chapter 2, then the study of the migration of a well-circulated research tool becomes a study of the trajectory through which the tool augments its epistemic capacity. For indeed, after making the transition from one discipline to another, the research tool in question gains new recognition for its expanded usages in the enterprise of knowledge production.

All that said, research tools do not simply work magically in multiple disciplinary contexts; instead, much work has to be done for such a transition to occur and to occur smoothly. In the case of formal language theory, it can be speculated that the ways in which the theory is used in aforementioned disciplines differ significantly as the users adapt the theory to suit different subject-matters and to meet particular disciplinary needs. As such, the term ‘tool migration’ aims to capture both the ‘situatedness’ of a research tool that was established in its native or home discipline and the effort it takes to ‘re-situate’ the tool in a foreign discipline. In particular, during the process of uprooting a research tool, significant contextual details may be stripped away, such as the implicit expertise or background assumptions that go into formulating a problem or interpreting the result of a calculation. Similarly, during the re-situation, new expertise and background assumptions may be introduced so that members of the importing discipline may use the tool to formulate a problem and interpret the results of calculation in a new context. Consequently, whatever might justify a tool's usefulness in the first place might not be carried over to the new context in which it is re-situated.

In addition to losing what makes the tool useful in the first place, there might be the risk of error particularly associated with using a research tool that is newly migrated into the discipline. Such errors include mischaracterization (i.e., characterizing the phenomenon in question with inappropriate assumptions), misinterpretation, (i.e., using ill-fitted background contexts to interpret the result), and misjudgment (i.e., falsely rejecting or accepting a novel use of a migrating research tool). Among these errors, mischaracterizing a target phenomenon and misinterpreting a computational result could lead to further errors of explanation in the new discipline or errors of prediction about the target phenomenon. Each of these errors, including misjudgment, incurs various costs (e.g., waste of resources, opportunity costs). It is thus desirable to treat a migrated research tool with care.

Taken together, by studying the applications of FLT across disciplines, this dissertation aims to shed light on the features that give rise to its usefulness across various disciplines and, conversely, the limitations that may have inadvertently escaped the attention of both scientists and philosophers of science.

1.3 My Approach, Research Questions, and Emphases

This dissertation seeks to contrast the role of the theory of formal languages in applications in three disciplines: linguistics, computer science, and more recently cognitive biology. My approach to this inquiry, however, is not meant to be historical. Instead, what I aim to deliver is a philosophical reflection on the epistemic issues related to its recent application in cognitive biology, i.e., drawing upon its 'reputation' in computer science to answer the question concerned with the 'species difference' between humans and other animals in terms of linguistic capacity. On the one hand, because FLT is at the intersection of multiple disciplines (it is especially regarded as a well-established component in the theory of computation in computer science), it is an attractive candidate for research applications. On the other hand, for the same reason, it is far from clear what renders the theory suitable for applications outside the disciplines of its inception and early development, especially the role it plays in cognitive biology — bridging between disciplines that are related to or in direct contact with the study of the mind and brain of both humans and other animals. Specifically, my research questions are:

1. What are the epistemic features that make it attractive to researchers from different, and not directly related, scientific contexts?
2. What might be the risks involved in producing disciplinary knowledge with the aid of a migrated tool such as formal language theory?

Each of these two questions concerns, respectively, the positive aspect and the negative aspect of the cross-disciplinary migrations of formal language theory. However, instead of treating these questions with equal attention, I put a slight emphasis on the second, i.e., the epistemic risks involved in tool migration.

Despite the recent interest to the topic of modeling and model transfer, the negative aspect of tool migration has not had the due attention they deserve from the philosophers of science. The discussion of model transfer concerns a relatively small set of mathematical models that are applied in multiple disciplinary contexts. One of the first philosophers of science to provide an account of model transfer, Paul Humphreys (2002, 2004, 2018) argues that versatile models are used to study phenomena of different domains because of the computational tractability they afford. Other philosophers (Knuuttila and Loettgers 2014, 2016) suggest that some versatile models provide not only computational tractability but also conceptual frameworks for theorization, which they label 'model templates.' These authors' analyses offer valuable insights, some of which converge with my finding (as I report in Chapter 3). At the same time, their unanimous omission on the risks inherent in this aspect of scientific practice calls for a remedy.

By contrast, in recent work that focuses on the use of evidence across disciplines, Mary Morgan (2010) presents a cogent analysis of why using evidence that is outside its site of discovery or construction, which she calls 'traveling facts,' needs to proceed with care. Specifically, facts could be received and treated as evidence in different ways according to the disciplinary base; once leaving the base where they were first found, facts might lose their usefulness and reliability. I take Morgan's discussion on traveling facts to be a starting point in my investigation. While my concern lies mainly in formal constructs instead of evidence (and much less about facts), as the chapters that follow will show, examining the usefulness of a research tool across disciplines prove to yield a fruitful result.

1.4 Organization

I organize the rest of this dissertation as follows.

In Chapter 2, I make two assumptions about the use of mathematical constructs in science. The first assumption is concerned with the instrumental characteristics of mathematical constructs, which allows me to view them as ‘research tools.’ I use game theory as an example and lay out three instrumental characteristics. Mathematical constructs are used to:

(1) formulate a research question into a tractable form,

(2) obtain an analytic result of the formulation, and

(3) interpret the analytic result as a solution or answer to the research question.

The second assumption is concerned with the contextual details of research tools. I suggest that research tools are ‘specific’ to a subject matter and ‘sensitive’ to the disciplinary context wherein it was first conceived. From these two assumptions (i.e., the ‘research tool assumption’ and the ‘subject-specific assumption,’ respectively), I argue that tool migration can be epistemically risky. Crucial disciplinary contexts can be lost in the ‘uprooting’ of a research tool or in the ‘re-situation’ of the tool, or both. One potential consequence is that whatever that makes the tool useful and reliable in the previous context fails to migrate with the tool into the new context.

To pursue this line of thought, I model my investigation of tool migration after Morgan’s (2010) work on ‘traveling facts.’ Morgan’s primary concern is the use of evidence outside its initial site of construction or discovery; her analysis of traveling facts thus reveals many features that the tools in migration also display. These common features—their role in knowledge production, their locality in origin, and their autonomy—prompt me to consider whether ‘losing integrity’ during tool migration hurts the cause of the scientific enterprise as it does in the context of traveling facts. However, despite rich in many ways, Morgan’s analysis of traveling facts lacks a criterion with which to systematically determine ‘integrity.’ According to Morgan, ‘losing integrity’ refers to a change in the content of a fact due to traveling. In my work, ‘losing integrity’ refers to a change in the content of a research tool due to migration. The difference between my work and Morgan’s lies in that I develop a clear view of what this ‘content’ would be when it comes to the tools in migration. Specifically, I determine the content of a research tool by its instrumental characteristics. The content of a research tool includes (1) the definitions of the theoretical terms that factor into formulating a research question and (2) the ways in which the analytic results are interpreted as an answer or a solution to the formulated question. For brevity, I refer to (1) and (2) as a tool’s ‘target profile’ and ‘usage profile,’ respectively.

In Chapter 3, I develop a four-fold typology as an ‘integrity’ check, which itself is based on two concept-pairs. First, between applications of a tool in migration, I distinguish an ‘established’ application and a ‘novel’ application. Second, within an application of a tool, I distinguish the tool’s target profile and its usage profile. Based on the changes (or the lack of changes) in these profiles between the established and the novel applications, I lay out four types of tool migration. In particular, tool migration refers to the practice of applying a research tool to study a subject matter that is ‘new to the tool’ (i.e., in so far as this tool’s previous applications are concerned). I analyze this practice into one of the four types as follow. Between the established and the novel applications:

- a ‘tool-application’ occurs when both the target profile and the usage profile of the tool remain similar,
- a ‘tool-adaptation’ occurs when the tool’s *target profile remains critically similar*,

whereas its usage profile has changed,

- a ‘tool-transfer’ occurs when the tool’s *usage profile remains critically similar*, whereas its target profile has changed, and finally
- a ‘tool-transformation’ occurs when both the target profile and the usage profile of the tool have changed.

To demonstrate, I examine an argument which suggests that losing integrity in tool migration is suboptimal. Using the migration stories of game theory as examples, I arrive at a negative conclusion; losing integrity in tool migration could, in fact, be optimal. Game theory originated in the study of mathematical models of strategic interactions between decision-making agents (von Neumann 1928). According to my analysis, its migration from the social sciences to biology and its homecoming from biology back to the social sciences shows two examples of tool-transformation (c.f. Grüne-Yanoff 2011a, 2016). Based on the analysis in Chapter 3, I argue that while ‘uprooting’ and ‘re-situating’ indeed undermine the tool’s integrity and while an exogenous research tool should be handled with extra care, losing integrity is not the problem. Losing integrity can be desirable and productive if it is resulting from proactive modifications to the tool as the users re-situate it into their respective disciplinary context. I conclude this chapter by discussing three concepts of epistemic risks associated with tool migration. For instance, modifying the target profile may help to prevent other users from the error of (1) mischaracterizing their target system with ill-fitting assumptions and (2) misinterpreting their analytic results. Most importantly, I pointed out that the inventor does not always have the best judgment of the limits of the tool that he or she constructed. I call this kind of error ‘misjudgment,’ i.e., (3) falsely rejecting or accepting a novel use of a migrating research tool.

In Chapter 4, I launch my investigation of the migration of FLT. I begin by reviewing two applications, one in linguistics another in computer science. Based these two applications, I argue that one crucial usage of FLT in linguistics is to theorize what kind of information processing device is best to describe the syntax of human language. By contrast, in computer science, the theory is not only applied to theorize about computability but also applied to engineer computers based on design needs. Thus, the migration from linguistics to computer science has augmented the usage profile of FLT. However, in contrast to the augmented usage profile, the application of FLT in computer science shows a target profile that is critically similar to its application in linguistics. Consequently, according to my typology, this particular migration can be categorized as tool-adaptation.

In Chapter 5, I explore one positive impact brought by applying the theory of formal languages in cognitive biology. Cognitive biology is the study of cognition as a biological function, one strand of which aims to understand both the evolution of and neural substrate for human linguistic capacity (Fitch 2014). It does so by applying FLT in experiments to test the abilities in nonhuman animals to learn artificial grammars. The first example of this approach is reported in Fitch and Hauser's (2004).

Bringing animal model organisms in contact with formal models of languages to investigate human linguistic capacity presents an intriguing and positive aspect of tool migration. On the one hand, using animal model organisms for the inquiry of human psychology has been a trademark of the behaviorist program, whose decline is often credited to Chomsky (1959a). On the other hand, the return of songbirds and nonhuman primates to the quest of understanding the faculty of human language is based on the classification scheme — the skeleton of which was single-handedly constructed by Chomsky (1956, 1959b). One may ask: How do scientists marry the behaviorist and the cognitivist approaches by bringing animals as model organisms back to the study of the linguistic capacity of humans? I argue that the key to this integrating branch of cognitive biology lies in an innovative insight from the migrating trajectory of FLT. In particular, the cognitive biologists 'enable' the Chomsky hierarchy—specifically, the scheme that computer scientists use to classify automata—to classify different systems of animal cognition, humans included. This innovation thus integrates animal model organisms with formal models of languages into one scientific enterprise. For this reason, I refer to Fitch and Hauser's innovative application 'the experiment of artificial grammar learning ('AGL') powered by the theory of formal languages' or 'FLT-powered-AGL.'

Finally, in Chapter 6, I discuss two negative reactions to applying the FLT-powered-AGL experimental protocols within cognitive biology. The first negative reaction comes from alert neurolinguists (Honing and Zuidema 2014) who disagree with Fitch over the usefulness and reliability of FLT in neurolinguistic studies. The second negative reaction is a misunderstanding of the intended use of the protocols. Specifically, some commentators interpret Fitch and Hauser's experiment as a test for 'recursion' (i.e., the capacity of 'infinite use of finite means' or 'open-endedness'). In another paper, Hauser, Chomsky, and Fitch (2002) suggest that recursion is a feature which bears unique significance to human language. That paper and Fitch and Hauser's work (2004) together prompted a surge of experiment searching for recursion in nonhuman animals. Unfortunately, the FLT-powered-AGL experimental protocols do not test for recursion.

This misuse of the tool brought backlashes to Fitch and Hauser's work and cast a shadow on their achievement.

In one of their efforts to clear the confusion, Fitch (2010) argues that the lack of proper understanding of FLT on the users' part is the source that caused this unfortunate search of recursion. In contrast, I argue that a misleading passage in the co-authored paper by Hauser, Chomsky, and Fitch (2002) is also culpable. This unfortunate episode of tool migration highlights the responsibility of 'importers.' Being an importer of a migrant research tool is of great responsibility. For by introducing a research tool to a new context, one may be exposing other users under an assortment of epistemic risks. Mistaking the tool for doing something which it does not is but one example. There may be more.

Chapter 2

Basic Assumptions and Concepts for a Study of Tool Migration

I begin my investigation by developing a conceptual framework with which I analyze the migration of formal language theory ('FLT'). Because my analysis will be touching on multiple disciplinary contexts wherein the theory has been applied, I aim for the framework to be somewhat general, i.e., not specific to any given disciplinary context, for a systematic analysis. The goal of this chapter is to start developing this very framework.

In Section 2.1, I discuss two assumptions that I take to be the starting point to my investigation: First, a research tool is a mathematical construct used in science in a certain fashion which justifies viewing it as a research tool. Relevant to the idea of a research tool are three instrumental characteristics: (i) formulating a research question into a tractable form, (ii) obtaining an analytic result from the formulated form, and (iii) interpreting the result as a solution or an answer to the initial research question. Second, I contend that the construction of research tools is both specific to the subject-matter and sensitive to the disciplinary context in which the construction takes place. In Section 2.2, I elaborate on what I mean by 'tool migration' and discuss why tool migration can be epistemically risky. In Section 2.3, I describe Mary Morgan's (2010) work on the cross-disciplinary use of evidence, which she calls 'traveling facts.' In Section 2.4, I recount her argument for why one should be cautious when using facts that have traveled away from the site of initial discovery or construction. Crucial to her argument is the idea of the ‘integrity’ of the traveling facts. A traveling fact traveled with integrity when its content remains “more or less intact during its travel” (Morgan, 2010, 12). Conversely, ‘losing integrity’ refers to a traveling fact that “changes so much during its travels that it is not recognisable as the same fact or has lost its credibility as a fact” (ibid.). Using her cautionary tale as inspiration, I continue in the next chapter to develop my analysis, essentially a typology with which to detect ‘integrity’ of a tool in migration.

2.1 Two Assumptions

A tool migration study makes at least two assumptions regarding its target of study. First, I assume that mathematical constructs used in science can be viewed as tools, and as I focus on the ways in which they are used in scientific research, I call them ‘research tools.’ Second, I assume that research tools are subject-specific, which means that each research tool is constructed for a certain subject matter. Let me elaborate on these two assumptions in turn. For brevity, I call them the ‘research tool assumption’ and the ‘subject-specific assumption.’ Throughout this chapter and the next chapter, I will use game theory—the study of mathematical models of strategic interactions between rational decision makers—as an example to illustrate these two assumptions. I reserve the discussion of formal language theory for Chapters 4-6 when I have sufficiently motivated the case and framework for studying tool migration. Also, the much more technical details required to introduce formal language theory prevents it from being an effective example in the early stage of building, as well as introducing, the framework.

2.1.1 The Research Tool Assumption

By a ‘research tool,’ I mean any formal construct in science that is to be *used* (or operated) in a seemingly systematic manner, and the outcome of whose operation is to be *interpreted* to answer a question related to a particular subject matter. Consider game theory. The operational aspect of a game theoretic analysis consists of three basic steps as follow.

First, to apply game theory, one needs to formulate a particular strategic interaction in game theoretic terms, such as ‘players,’ ‘acts’ and their ‘payoffs.’ This process involves identifying who the players are, what may be their acts, and assigning a numeric value to each of these acts as payoff before representing the interaction in a matrix. Second, having assigned the values to the acts, one calculates to obtain an analytic result of this formulation, such as the Nash equilibrium. Third, and finally, one interprets this above-mentioned result as the solution to the question regarding a particular strategic interaction. In other words, to apply game theory is to follow these three basic steps (‘problem formulation,’ ‘calculation,’ and ‘interpretation’).[3]

Moreover, concerning especially the interpretational aspect of a research tool, a user obtains a result after going through the steps of problem formulation and calculation. Note that this result can be properly understood only through the lens of a certain interpretation. For example, a Nash equilibrium is a set of acts or moves in which every agent will be better off maintaining his or her decision given other players doing the same. In a sense, the Nash equilibrium of a game-theoretic analysis is a meaningful

---

[3] My analysis does not exclude that in actual practice, one may need to go through these steps multiple times to fine-tune the result.

‘solution’ in virtue of the usual understanding of the game-theoretic formulation of a problem.

To put the analysis above in general terms, I contend that a research tool guides its users to

i. formulate a problem concerning a subject matter (‘problem formulation’),

ii. obtain an analytic result for the formulated problem using the tool (‘calculation’), and then

iii. interpret the result as a solution to the problem, which advances users' understanding of the subject matter (‘interpretation’).

Jointly, problem formulation, calculation, and interpretation give rise to the instrumental characteristics of formal constructs in science that justify viewing them as ‘research tools.’ To drive home the point, when applying such a construct, the goal of performing a sequence of prescribed steps as depicted in (***ii***) goes beyond merely completing the calculation to obtain a result. Instead, as described in (***iii***), this result itself is to be interpreted in a certain way so that one may solve a problem, answer a research question, or, eventually, gain knowledge about a subject matter. Indeed, manipulating symbols is a means to the end that was specified in (***i***) during the mathematical formulation of the initial problem. Thus, a formal construct that exhibits these above-mentioned characteristics in scientific practice can be seen as a research tool because it assists its users to meet an end.

To complete my sketch of a conceptual framework toward a tool migration analysis, I now turn to the second assumption: the subject-specific assumption.

2.1.2 The Subject-Specific Assumption

Like any product of scientific endeavor, research tools are constructed in a particular place and time. Some of them are further developed and eventually receive recognition as a viable tool within a certain scientific context. Assume that a scientific context is to be characterized by its subject matter, metaphysical commitments, methodological assumptions, and the accepted hypotheses or research questions. Then the recognition of a certain construct as a viable research tool is context-specific. That is, the recognition of a research tool being useful and reliable is granted given all the characteristics of the scientific context or tradition in which the tool is constructed or developed. In other words, a certain construct becomes an established research tool when it is acknowledged by its users in a particular scientific context. It follows that the very scientific context's characteristics determine the said tool's research capacities (e.g., to what subjects it may be justifiably applied or in what ways it could legitimately be used to produce knowledge).

Among the characteristics of a scientific tradition, the subject matter is arguably the most salient feature regarding a research tool. For example, the subject matter of game theory is strategic interaction between rational agents, individuals or groups. For example, the subject matter of game theory is strategic interaction between rational agents, individuals or groups. Research tools are subject-specific in the sense that they are constructed to carry out a series of actions such that certain questions regarding their subject matter can be answered.

2.2 What is Tool Migration and Why it Needs Attention

Having introduced the concepts of ‘research tools’ and the ‘subject-specific’ assumption, I now specify what I mean by a case of tool migration:

> When a formal construct, which has been recognized as a viable research tool for a particular subject matter in a particular (sub)disciplinary context in science, is allegedly used for different subject matter in an another (sub)disciplinary context, it is a case of tool migration.

Being subject-specific does not prevent a research tool from being used in investigations about other subject matters. Indeed, certain advancements in science were made by discovering novel uses of tools that had been constructed to study something else, such as the use of Faraday's mechanical model of fluid motion to mathematically characterize the electromagnetic field (Maxwell 1861). At the same time, the subject-specific assumption does entail that *when a tool is used in a novel way, some kind of justification is necessary*. Thus, when such a justification is missing or misled, the novel use of an established tool can raise epistemic concerns. For instance, game theory—developed to model strategic interaction between rational agents (von Neumann 1928)—has been borrowed to study biological evolution (e.g., Maynard-Smith and Price 1972). Formal theoretical frameworks, such as game theory, are appealing candidates of tool migration perhaps because they are abstract and thus thought to be more generally applicable than, say, the oxygen theory of combustion. However, on closer inspection of specific applications, e.g., game theory used in evolutionary biology, one may question whether game-theoretic notions and analyses are appropriate for generating knowledge about biological evolution—where assumptions about rational agents and their strategic reasoning may not apply. These concerns motivate this present dissertation that I term a study of tool migration. Because without some deeper understanding of what is behind the tool and what may justify its novel use, the status of the result of using an established research tool in a novel way is unclear at best.

In other words, the 'research tool' assumption and the 'subject-specific' assumption jointly suggest that tool migration can be epistemically risky. When an established research tool appears in use outside its discipline of origin, whether the tool's efficacy gets carried over becomes a question. Thus, cases of tool migration in science call for attention. In the literature, there have not been explicit discussions concerned with the epistemic risks associated with tool migration, except for Mary Morgan's cautionary tale for the use of traveling facts.

2.3 Be Cautious of 'Traveling Facts'

In "Traveling Facts" (2010), Morgan reports three features of facts that are germane to the caution of using the 'well-traveled' ones. First, facts are locally discovered or produced. Second, facts are used to generate further knowledge. Third, and, most crucial, facts are autonomous — that is, once released to the community, facts gain a life of their own; users may then find novel ways of using those facts, within or outside their initial context, with or without a blessing from the original producer or discoverer of the facts. Based on these observations, resulting as a summary of a four-year-long research project devoted to analyzing the dissemination of facts, Morgan argues that one should be cautious when using facts that have traveled ‘far’ and ‘wide,’ i.e., having been through multiple contexts that do not share much in common.

According to Morgan, a fact is a piece of knowledge, locally generated, accepted by the members of its local community for its usefulness and reliability. A traveling fact is one that finds new users, new uses, or that morphs into new shapes (e.g., becoming fiction) within or outside its local community. Before unpacking her argument, I should note that by ‘facts,’ Morgan does not merely refer to ‘true statements’ of some sort. To her, facts come in different forms. They “may be expressed in linguistic statements ...; they may appear in pictures, diagrams, models, maps, documents, biographies or novels; they may be found as material facts located in artefacts” or “as numerical constructions about the future of our overheated planet” (Morgan 2010, 8). More relevant to our present purpose, facts may also be “expressed in the behavioural characteristics of” organisms in an experimental setting or in “statistical and mathematical models” (ibid., pp. 8/27). In addition to being elements of knowledge, this wide range of things that Morgan calls facts, share the following features: they are usable, generated locally, accepted by their communities as useful and reliable, and finally autonomous.

Invoking the metaphor of traveling, Morgan (2010) focuses on characterizing the dissemination of evidence that is produced or discovered in a wide range of areas. To illustrate, I will use her examples from ethology, climate science, and the medical case report system. Morgan’s work on traveling facts, which I briefly review as follows, provides a point of departure for my analysis of tool migration in Chapter 3.

2.3.1 Facts are usable

In Morgan's (2010, 8) words, “facts are a usable category.” Some facts are used to produce additional knowledge. For instance, the medical reports of several young men in New York showing an unusual coalition of symptoms led to the recognition and characterization of a then-new disease, the HIV-AIDS syndrome.[4] Some facts are used to make decisions. For example, medical case reports of well-known, highly infectious, diseases such as measles or flu, are used, through statistical methods or simulations, to predict the spread of the diseases. In light of these predictions, public health authorities choose proper responses to better control the spread.[5] For another example, Niko Tinbergen and Konrad Lorenz reported in the presence of overhead moving silhouettes of predators, certain species of birds on the ground instinctively take cover. This ethological discovery has inspired the design of birds-of-prey window decals in silhouette style, intended to reduce the number of small birds flying into the glass.[6] Window decals of this sort have not only been available for purchase but also seen in the display at locations across continents. However, just because facts travel to be used elsewhere does not mean that they always travel well.

While Tinbergen and Lorenz's scientific discovery seems to have remained suitably qualified, those window decals turn out not to work as intended. According to Morgan’s report, genuinely flying birds do not seem to avoid stationary silhouettes of

---

[4] Ankeny 2010, summarized in Morgan 2010.

[5] Mansnerus 2010, summarized in Morgan 2010.

[6] Burkhardt 2010, summarized in Morgan 2010.

their predators. Nonetheless, the National Audubon Society—a non-profit organization for the conservation of birds, other wildlife, and healthy ecosystem—advises bird-lovers to use window decals to prevent bird collision. In the instructions, it is emphasized that the decals must be spaced “no more than 2-4 inches apart.”[7] One can argue that in that case, what essentially deters wild birds from colliding into the windows is the density of the decals, which creates a visible barrier to the birds. Thus, the shape of these window decals is largely irrelevant. In this example, Morgan suggests, we have seen a scientific discovery travel to inspire and influence decisions in the practical realm (i.e., the production and the consumption of these decals), only to find itself turned into some kind of urban legend. Such an ending for a traveling fact is undesirable and yet not exceptional as it could happen to facts generated in other areas. In some scientific domains, for instance climate science or epidemiology, successful traveling of facts can be crucial to the well-being of humans, which Morgan takes to be an important motivation for a proper understanding of traveling facts.

2.3.2 Facts are generated locally

No matter how far and widely facts venture away from their origins, they are typically “generated in a particular time and place and developed by particular individuals or groups of scientists” (Morgan 2014, 1013). To Morgan, facts are local in both the spatial-temporal sense and the sociological sense. Facts, regardless of whether they were discovered or produced, have their “place of origin” (Morgan 2010, 8). This point is demonstrated even in the medical case reports that became the early diagnoses of

---

[7] Audubon.org 2014 http://www.audubon.org/news/help-birds-avoid-deadly-collision

HIV-AIDS. It is true that the recognition of the new disease took several medical reports from different sources, but the documentation of each of those individual reports occurred in a particular location at a point in time. In terms of their sociological attribute, behind a discovery or a production of facts is always *a discoverer or a producer* who handed the fact to other members in the community. The sociological locality of a fact, thus, refers to the immediate community members who would become the first group of people to accept the discovery or production to be reliable and useful.

An important assumption in Morgan's account about this locality of facts is that the criterion of ‘facthood’ is community-dependent. That is, she does not rely on, nor does she attempt to develop, an objective criterion for picking out what may be facts. Instead, it is the behavior of a community toward a particular piece of discoveries or intellectual creations that marks facts from non-facts. On the one hand, according to Morgan (2010), facts are not to be understood “as an expression of that community's *belief*, or *opinion*, but rather that such a community has good reasons to take those things as facts, and will be likely to have the confidence to act upon them as facts” (p. 11, original emphases). A community is said to have come to ‘accept’ things in their community as facts when they are “sufficient for people to act upon them or use them in support of their action” (ibid., 11). On the other hand, being recognized as a fact in one community does not automatically make it so in another. The context that makes facts useful and reliable in the community does not travel as facts do; it stays in the community. At the same time, of course, the locality of facts does not constrain them to stay only within their immediate community. This leads to the autonomy of facts that I discuss next.

2.3.3 Facts are autonomous

Facts, Morgan writes, “turn out to be like children: Their parents who found or fashioned them soon lose control of them, they leave home, their product markings become lost as they make their way into all sorts of other unknown communities and fulfill all sorts of unexpected purposes” (2010, 26). This metaphor captures “the tendency of facts to travel relatively independently” without “much reference to the context in which they were produced” (ibid., 26).

Facts leave their place of origin and immediate community to be used elsewhere, but not all of them travel well. We have seen the example in which an ecological discovery turned into some kind of urban legend. According to Morgan (2010), facts concerned with climate change provide another example.

> We all know about climate change from the scientists, but these facts did not travel easily to us. We all know now that the world’s climate is getting warmer, but for a long time, we were not very sure what facts we knew: how certain it was, how serious it was, how fast the change was happening, how different bits of evidence fitted together to form a consistent account and how far different scientists were in agreement about it. And we still don’t know much about how it will affect different parts of the globe. (Morgan, 2010, 4)

In turn, these facts about climate change did not enter the practical realm soon enough where they could make the (potentially much) needed impact. Successful travels of facts can be crucial to the well-being of humans. Thus, Morgan asks: What makes facts travel well so that the dissemination of them brings positive influence instead of misinformation? In order to answer that question, she first addresses what it means for facts to travel well.

2.3.4 Traveling fruitfully and traveling with integrity

Morgan begins her analysis by examining two ideas: A fact travels well when it travels fruitfully and when it travels with integrity. 'Traveling fruitfully' refers to the fecundity of the journey:

> [F]acts may travel far and wide in terms of time and of geographical and disciplinary space to find new *users*. More unusually, perhaps, are the ways in which facts find new *uses*: They gain new functions, coalesce in new patterns and make new narratives. (2010, 12 emphasis original)

In other words, the increase of new users, new uses, or new communities, suggests a well-traveled fact. This idea of 'traveling fruitfully' is meant to capture the intuition, as Morgan contends, that facts that encounter resistance outside their immediate community cannot be said to have traveled well. For instance, the number of 'climate change skeptics' indicates that facts that are established in climate science are thought to be fictitious by some communities.

By contrast, 'traveled with integrity' refers to the stability in the content of the traveling fact, namely, whether the content remains "more or less intact during its travel" (2010, 12). As Morgan puts it, "if a fact changes so much during its travels that it is not recognisable as the same fact or has lost its credibility as a fact, it would be hard to claim that the fact has travelled well" (ibid.). Taken these two aspects together, facts that travel well are those that expand their influence in the dimensions of *users*, *uses*, and *communities*, all the while without compromising their content.

2.3.5 Adventurous travel of facts

It turns out that, according to Morgan (2010), if one accepts the community view of facts, there is a potential trade-off between traveling fruitfully and traveling with integrity. Consider the justification based on which each community accepts or denies particular things as facts. Justifications of this sort are derived from the "integrated clusters of knowledge elements" previously established within a community (Morgan 2010, 14). For instance, facts "often have details that we might call qualifications," which are "circumstantial or contextual elements that contribute to their usefulness (ibid., 16). These qualifications that may be well-understood but implicit within a given community are thus an essential component of the 'facthood' of a fact within that community. However, when facts travel, these qualifications, whose relevance to the acceptance and meaning of the fact, might not be evident to those outside the community.

Moreover, facts may "pick up extra elements on their travels and become covered with additional elements or even sharpened in certain ways" (ibid., 16). It follows, Morgan argues, that leaving its community of origin "may subtly change the nature" of the traveling fact (ibid., 15). Thus, the further the facts travel to meet new users in new communities, the more challenging it becomes to maintain the integrity of their content.

The situation worsens once we consider seriously the possibility that the appropriate use of some facts requires expertise that is community-specific. On the one hand, Morgan says, "[a]t its most fruitful, the use of travelling facts creates a new pattern, a new coherence, a new narrative or fulfills a new role" (ibid., 20). On the other hand, "[t]hese more adventurous travels of facts, and their ostensible contrasts, have to be carefully looked out for" because "[t]oo often, what appears to be freely travelling knowledge is, in fact, dependent on tacit or expert community knowledge to make it transfer effectively" (ibid., 26). That is, she continues, "such knowledge is not in general separable from its base in techniques and expertise" (ibid., 26). As a result of the potential trade-off between traveling fruitfully and traveling with integrity, one should be cautious about well-traveled facts.

Nonetheless, despite being local in origin and community-dependent in their usefulness and reliability, some facts (as she argues) have traveled well, so the question becomes: What makes facts travel well? Morgan suggests that the success of a fact in a community does not determine, nor predict, whether a different community will consider its content useful and reliable. Instead, she argues, what makes facts traveling well is the "users, in different times, places and disciplines, with different questions and different purposes, who largely determine the uses of facts at various destinations, and thus how well they have travelled to fulfill new purposes" (ibid., 25).

Eventually, Morgan concludes that good travel companions are crucial. In order for traveling facts to be appropriately acted upon, or acted with, in the hands of new users in new contexts, it takes competent curators to provide correct labeling and proper packaging of the facts at the point of introduction to a new community. Morgan goes on to elaborate the labeling and packaging practice using different cases of traveling facts, a discussion full of insight on its own right, but I will now turn to appraise the concept of 'traveling with integrity' in Morgan's account. For a concept playing a critical role in judging whether a fact has traveled well, my assessment of it, in particular the concept of ‘losing integrity,’ shows room for improvement.

2.4 A criterion of integrity for traveling facts? A critical assessment

There are two reasons why a criterion of integrity is important to Morgan’s treatment of traveling facts. First, the goal of Morgan’s project was to extract information from successful cases what makes facts travel well, i.e., facts that manage to travel both fruitfully and with integrity. For this approach to succeed, one needs to be able to accurately distinguish the facts that have traveled well from those that have not. Second, according to Morgan, due to the threat of ‘losing integrity,’ one should be cautious when using facts that have traveled through multiple contexts. Her argument goes roughly as follows:

1. The components that are used in the production of knowledge within a particular community are facts to that community in the sense that they are accepted by that community as being useful and reliable.
2. Such acceptance is based on a cluster of preexisting knowledge, qualifications and background assumptions, which may or may not apply to the context of other communities.
3. Traveling facts are facts circulating across communities.
4. However, traveling facts do not travel with the qualifications and background assumptions that granted their usefulness and reliability in the first place, a phenomenon that she calls ‘losing the integrity.’
5. As a result of detaching from its base, ‘traveling’ undermines the integrity of the facts, such that one should be wary of using traveling facts as they may no longer

> be as useful and reliable as they were in their home community.

Putting together, the concept of integrity plays a pivotal role in Morgan's analysis, justifying caution whenever traveling facts are in use.

While Morgan is not concerned with what I call research tools, her evaluation of why one should be careful when using traveling facts is informative to my investigation of tool migration. It is possible that users of research tools in migration should exercise care, and it is also possible that what justifies such caution is due to 'losing integrity.'

Before I proceed to explore these two possibilities in Chapter 3, it is crucial to take note of how Morgan determine the integrity of traveling facts, a concept I will need to develop for my inquiry.

To Morgan, it is inevitable that facts change some aspect or another due to travel, but not all such changes necessarily threaten to undermine their integrity. In her words:

> [T]ravelling facts as rubber balls: They have a certain shape; they can be carried, rolled, squeezed, bounced, kicked and thrown without harm to them; and they can be used in many different ways and in different situations. (2010, pp. 15-6)

What would it be like, one asks, for a fact to have changed in 'shape' like a rubber ball without harming the integrity of its content? According to her, despite superficial changes, the integrity of the content of a traveling fact can be recognized as we see it. Here is an example Morgan (2010, 16) offers to illustrate the possibility of a fact traveled far and wide without having undermined its integrity:

> A good example is found in the travels of the classical style of architecture from ancient Greece to nineteenth-century America. Certain details of the style were altered in the process of adaptation, such as the nature of the materials, the

> addition of windows and the reversal of light and shade in the exterior. While the community of American architects and builders seemed to delight in their ability to adapt the stylistic facts of classical architecture to their own contexts, they also observed certain boundaries. For despite their additions, alterations and subtractions, there is no doubt that such buildings retained a recognisable integrity as examples of "the classical style" in their new domain (Schneider [2011]). This is what we mean when we suggest that facts that travel well exhibit a strong degree of integrity, but they also have a degree of squishiness, a squishiness that may result in them getting their hard edges rubbed off, changing their surface elements or gaining some additional covering as they travel.

Morgan was not being figurative when speaking of facts changing in shape. The "additions, alterations and subtractions" effectively changed the implementation details of an architectural style. However, the style survived, as she suggests, because one can still recognize it being the same style (Morgan 2010, 16). Essentially, Morgan's approach to the integrity of traveling fact is that of 'we know it (i.e., whether it is the same fact) when we see it.'

Morgan's approach to integrity in the context of traveling facts raises more questions than it answers. For instance, her focus is the discovery, production, and dissemination of evidence, or what is sometimes thought of as hard facts. These facts include "modern statistical and mathematical models" (2010, 27) along with medical records and artifacts. On the one hand, it is understandable that she does not attempt to describe how one may systematically determine the integrity of these varying types of facts. On the other hand, without a generally applicable criterion for the integrity of facts, it is not clear how her conviction to the possibility of successful travel of facts applies to facts of types other than the architectural styles. More importantly, when the way with which to select ‘successful examples of traveling facts’ lacks transparency, the extracted information of how facts travel well is either limited in applicability at best or misleading at worst. For instance, consider three different communities, A, B, and C. By following the method of ‘we know it when we see it,’ a given member in Community C may not have the competence to discern the integrity of a fact that travels from Community A to Community B. Morgan’s approach is thus opaque to the out-group members who do not have necessary information of, but need to decide on the matter regardless.

What I take from Morgan’s discussion is an idea for improvement: it is important for a study of tool migration to establish a criterion of integrity in its analytic framework. The subject-specific assumption about research tools, which I subscribe to, has a similar effect on tool migration as does Morgan’s community-specific assumption about facts on traveling facts. Research tools are constructed for different subject-matters within different disciplinary contexts. One might charitably assume that the members of a given discipline have good reasons for gravitating towards a few tools from the pool of all tools constructed within his or her discipline. In contrast, making the same assumption about their handling of exogenous tools seems dubious. All that said, one of the goals of my dissertation is to decide what may be the source of epistemic risks in using a tool that has migrated. ‘Losing the integrity’ thus makes a reasonable point of departure, which I turn to in Chapter 3.

Chapter 3

Be Cautious of Migrant Tools? Yes, but Losing Integrity is not to Blame

In this chapter, I explore whether losing integrity is indeed suboptimal in the context of tool migration. By studying the migration of game theory as an example, the answer I eventually arrive is negative. In Section 3.1, I draw a parallel between traveling facts and migrating tools in terms of their epistemic features, such as their role in knowledge production, their locality in origin, and their autonomy to be used beyond their origin. In Section 3.2, based on these features, I pose an argument for examination. This argument urges caution about using research tools that have migrated. This argument closely follows Mary Morgan’s (2010) analysis of losing integrity in traveling facts. It argues that because there are many features in common between traveling facts and the migrant research tool, one may speculate that losing integrity undermines successfully applying a migrant tool as it does to the applications of traveling facts. In Section 3.3, I develop a criterion of the integrity of research tools with which to analyze the migration of game theory. Unexpectedly, my analysis, which I discuss in Section 3.4, suggests the opposite of what I have speculated. In Section 3.5, I conclude my observations of the migration of game theory. Through studying this example, one learns that proactively modifying a migrant tool helps the users to avoid epistemic risks such as mischaracterizing the target phenomenon with an ill-fitting profile. Because modifying a tool entails altering the content of the tool, which in turn entails ‘losing integrity,’ I argue that although one should remain cautious of using a migrant research tool, losing integrity is not to blame. I end this chapter by a discussion of two further questions about tool migration, which I explore in Chapters 4 and 5.

3.1 Common Features between Traveling Facts and Research Tools

Research tools may not be what Morgan thinks of as ‘facts,’ but so long as they exhibit features sufficient for issuing caution of traveling facts, the use of exogenous tools deserves attention. In Morgan’s (2010) discussion, these features include ‘being used to generate further knowledge,’ ‘being locally generated,’ and ‘being autonomous.’ Together, these features give rise to a potential trade-off between accumulating more users or uses and maintaining the integrity of the content carried in the traveling fact. I will talk about each of the three features in the context of tool migration in succession.

3.1.1 Research tools are useable

Research tools are usable in advancing knowledge. Previously in Section 2.1, I introduced the concept of ‘research tool’ to analyze formal constructs in science that are used to advance the user’s understanding of a subject matter. I argued that formal constructs used in such a way may be viewed as research tools because they serve as a means to an end. Merely completing the calculation to obtain a result is not the end because the result is to be interpreted in a way so that it becomes an answer to the question, or a solution to the problem, formulated in the first place. In other words, one gains knowledge of a subject matter by manipulating symbols as a means to the end that was specified during the mathematization of the initial problem. Thus, it is entailed in my definition that research tools are usable and used in producing further knowledge — as is the case with facts. Of course, instead of the question ‘whether’ research tools are used to generate further knowledge, a more pressing question to my inquiry is ‘how’ they are used to do so.

Recall Section 2.1. I proposed to analyze an application of a formal construct in science in terms of ‘problem formulation,’ ‘calculation,’ and ‘interpretation.’ These three steps of using a formal construct jointly give rise the construct an instrumental character. In that chapter, I used game theory as an example to illustrate. Here, with another example, I would like to highlight the importance of a *background context* when applying a formal construct to advance knowledge. Consider the formula for the population mean:

$$\mu = \frac{\sum_{i-1}^{N} x_i}{N} \tag{3.1}$$

The symbol ‘μ’ represents the average score of a population on some variable, which can be obtained by completing the calculation expressed on the right-hand side of the equation. To do so, a user will first identify the score on some variable of each individual in the population. The user will then take the sum of all scores presented in the population (i.e., x_1, x_2, x_3, ...) and obtain μ through dividing the sum by the total number of individuals in the population (i.e., N). Note that, the value μ as a result of the above calculation is informative only in light of some context. For example, the individual scores may be the SAT scores, and the population may be the newly admitted applicants at a certain university. However, without being given such a context, even to the user who completed the calculation, the resulting μ is not more meaningful than, say, any value one looks up in the multiplication table. Moreover, without comparison of some sorts (e.g., with the national average or with the average of the previous year's record), the admission statistics of a single college in a given year provides limited understanding. Thus, the formula in (3.1) is useable to generate knowledge partly because of the analytic procedure it provides (i.e., by which a user calculates to obtain a value as a result) and partly because of the connections that the user is making between the resulting value and a particular background context. Such a context is usually given at the stage of formulating the problem. Taken all together, it is with the context that users of an equation, or any formal constructs, may meaningfully interpret the result of his or her calculation.

3.1.2 Research tools are local in its origin

Research tools are, figuratively speaking, local in origin. The locality of a research tool concerns both the *construction* of the tool and the *acceptance* of it within a scientific community. Regarding the construction aspect, research tools typically start with a particular initial setup and question. The initial target system plays a role in the mathematicalization of the phenomenon of interest, which in turn allows mathematical proofs to be found. Game theory, for example, is local insofar as it originated in the mathematical formulations and theorizations of living-room games, e.g., roulette, chess, bridge (von Neumann 1928). John von Neumann (1959, 13, translated from 1928 pp. 295-320 by Sonya Bargmann) begins his article that is now considered the origin of game theory by posing the question as follows:

> n players S_1, S_2, …, S_n are playing a given game of strategy, *B*. How must one of

> the participants, S_m, play in order to achieve a most advantageous result?

He then continues to elaborate on what might be a ‘game of strategy’ (1959, 13, emphasis mine):

> A great many different things come under this heading, anything from roulette to chess, from baccarat to bridge. And after all, any event — given the external conditions and the participants in the situation (provided the latter are acting of their own *free will*) — may be regarded as a game of strategy of one looks at the effect it has on the participants.

For instance, in the passage where von Neumann (1959, 14, emphasis mine) derives what he calls a qualitative description of the game concept, he writes:

> We may assume that ... [a] game of strategy consists of a certain series of events each of which may have a finite number of distinct results. In some cases, the outcome depends on chance, i.e., the probabilities with which each of the possible results will occur are known, but nobody can influence them. All other events depend on the *free decision of the player* S_1, S_2, …, S_n. In other words, for each of these events it is known which player, S_m, determines its outcome and what is his state of information with respect to the results of other (“earlier”) events at the time when he makes his decision. Eventually, after the outcome of all events is known, one can calculate according to a fixed rule what payments the players S_1, S_2, …, S_n must make to each other.

By formulating this above discussion of a *game of strategy*, von Neumann derives five “rules of the game” (ibid, 14), which give him the bases to prove the minimax theorem for two-person zero-sum games (Kjeldsen 2001).[8]

Additionally, the acceptance of research tools has a local component. Consider game theory again. Certain assumptions (e.g., regarding the ability of the players to make choices in the game, the players' rationality and knowledge about the game and about other players' rationality and knowledge about the game and so on) are crucial to the tool's initial applications in economics (von Neumann and Morgenstern 1944). However, disciplines that do not readily accept these assumptions about their human subjects or subject matter (e.g., cultural anthropology) may not employ game theory in their chain of knowledge production. Thus, both the constructing of assumptions and the initial success of a research tool are specific to the disciplinary context, and hence both are local in origin.

3.1.3 Research tools are autonomous

Despite their locality, research tools are autonomous; upon being released, their inventors lose control over them. Users from within or outside the native discipline of these tools (i.e., where they were initially conceived) may freely apply them in research *with or without* the inventors' blessing.

For instance, John Maynard-Smith took inspiration from game theory to develop an account of biological evolution (as Grüne-Yanoff argues [2011a]). The resulting tool, including the analysis of evolutionarily stable strategy ('ESS'), is now known as evolutionary game theory ('EGT'). This tool (EGT) has subsequently migrated back to economics, where game theory was originally established. However, what may not be

---

[8] The history and evolution of different versions this proof is reported in Kjeldsen (2001).

widely known is Maynard-Smith’s blatant disapproval of EGT’s homecoming migration. In an interview (Maynard-Smith 1997), he publicly criticized the use of EGT in the context of theoretical economics:

> I'm now amused to discover that the economists and so on are trying to borrow evolutionary game theory back and introduce it into economics. I'm very unconvinced that they're justified in doing this — because evolutionary game theory hinges on the notion of heredity, i.e. the essential notion being that your success in the game determines how many children you have, and your children are like you, and the whole thing hinges upon that essential assumption. And I can't see that that assumption really holds in economics.

Maynard-Smith’s critique of EGT's homecoming requires closer scrutiny. In fact, I will come back to this episode later in Section 3.4 and suggest that it represents an example of what I call a ‘misjudgment’ related to tool migration. For now, let me turn to explain how these above features may result in precarious tool migration.

3.2 Framing the Inquiry

In what follows, I adopt Morgan's analysis of traveling facts to consider research tools in migration. To begin, let us think of tool migration in terms of ‘uprooting’ and ‘re-situating.’ In our present concern, ‘uprooting’ captures the sense in which a formal construct is taken from its base where it has been established as viable for generating further knowledge. ‘Re-situating,’ in contrast, captures the effort of putting the formal construct to work in a novel context where it is expected to perform a new task, such as answering a new research question. Thus, similar to traveling facts, research tools could lose a host of qualifications and background assumptions that granted their usefulness and reliability in the first place. Such an incident could occur either in the process of uprooting or re-situating, or both. For further investigation, I organize these above observations into three statements as follow:

a) Due to 'uprooting' and 're-situating,' migrating a research tool from its native discipline to use in a novel discipline may undermine the tool's 'integrity' (i.e., tool migration gives rise to losing the integrity of the tool).

b) Losing integrity compromises the usefulness and reliability of applying a tool that has migrated (i.e., losing the integrity of an exogenous tool is counterproductive and, hence, undesirable).

c) Therefore, an exogenous research tool should be handled with extra care.

Statement (***a***) contends that tool migration may lead to losing the integrity of the tool in migration. Moreover, Statement (***b***) suggests that losing integrity in this sense may lead to unsuccessful applications of the migrating research tool because the tool's usefulness and reliability would be compromised. Finally, Statement (***c***), in the form of a conclusion from the previous two statements, urges care when using a research tool outside its native or home discipline.

In the sections that follow, I argue that when both Statement (***a***) and Statement (***c***) are correct (i.e., a research tool may indeed lose its integrity due to migration and one should handle migrated tools with care), Statement (***b***) can be false (i.e., losing integrity is not the source of problem). In particular, I will use the migration of game theory as an example and show that while one should cautiously apply a migrated tool in novel contexts, what licenses wariness is not 'losing integrity.' Instead, proactively modifying the tool, which inevitably undermines the tool's integrity, is crucial to successfully apply a migrated tool in novel contexts. My argument relies substantially on a criterion for determining the integrity of a research tool, which I now turn to in the next section.

3.3 A Typology of Tool Migration as an 'Integrity Check'

The 'integrity' of a migrating research tool, as I use it, refers to the tool's state of being whole or unmodified. To detect a potential loss of integrity, one may need to perform an 'integrity check.' The stress on an integrity check is the point where my approach departs from Morgan's. As mentioned in Chapter 2, Morgan did not attempt to provide a systematic approach to the integrity of traveling facts. Unlike Morgan, I do not have the liberty to avoid doing so — my argument hinges on being able to state clearly in what sense migration undermines the integrity of a research tool, especially because the goal of the present investigation is to decide whether losing integrity is indeed the reason why one should be cautious of using a migrated research tool. To this goal, I develop an 'integrity check' based on the instrumental characteristics of formal constructs that I analyzed in Section 3.1. In particular, I single out two major kinds of changes within a research tool that are significant for the tool's performance in knowledge production.

The first kind pertains to 'problem formulation' and considers all the definitions of the theoretical terms and the relation between them in the formal construct. For instance, game theory contains definitions of a 'game,' its 'players,' 'an act,' 'payoff,' as well as how these terms relate to one another. For simplicity, I call the collection of theoretical terms and the relation between them the tool's 'target profile.' Essentially, a target profile of a tool is the collection of theoretical terms allowing users to 'profile' the phenomenon he or she targets.

The second kind pertains to the 'interpretation' aspect of applying a research tool. For instance, in a game-theoretic analysis, a Nash equilibrium could be understood differently depending on what the initial problem was and, consequently, what was needed to solve the problem. Concerning a strategic interaction between competing firms or individuals, a Nash equilibrium could be interpreted as an explanation or a prediction regarding the outcome of the interaction. Alternatively, it could be used to optimize the result of strategic interaction. I call the collection of the ways in the result is interpreted, e.g., as a description, a prediction, a way for optimization, or an explanation of a given aspect of a target phenomenon, the tool's 'usage profile.' To summarize:

- Target profile: all definitions and relations concerning the theoretical terms of the formal construct in question
- Usage profile: all the ways in which the calculation result of the tool is interpreted (i.e., all the ways in which the tool is intended to be used)

Moreover, in order to trace the trajectory of a migrating research tool, I distinguish between two applications of a tool the 'established use' and the 'novel use' of the tool:

- Established use: a representative application of a research tool in its home discipline
- Novel use: a recent application of a research tool in a novel discipline

With these two concept pairs, I classify four different types of tool migration, based primarily on the differences or similarities between the established application and the novel application of the migrating tool in question. For instance, when a tool retains both its target profile and usage profile in the novel use, that particular use of the tool is considerably similar to its previous application; I call such a case **tool-application**. In contrast, when both the target profile and usage profile have changed in the novel use, the migration is transformative to the tool, and therefore I call it a **tool-transformation**. Between these two extreme types, there may be cases where the novel use alters only one of the two profiles but not both. Thus, when a novel application changes a tool's target profile but not the usage profile, I call it a **tool-transfer**, whereas when a novel application changes the tool's usage profile but not the target profile, I call it a **tool-adaptation**. See Table 3.1 for a summary.

Table 3.1: A Typology of Tool Migration
Four potential relations between an established use and a novel use of a research tool

		USAGE PROFILE	
		REMAINS	**DEVIATES**
TARGET PROFILE	**REMAINS**	Tool-application	Tool-adaptation
	DEVIATES	Tool-transfer	Tool-transformation

As a proof-of-concept, in the sections that follow, I apply the above typology to analyze the migration of game theory. Relating to the argument that I set out to examine, which regards losing integrity as suboptimal in the context of tool migration, my analysis will emphasize two major points. First, the migration of game theory from the social sciences to biology can be viewed as a case of tool migration. As I argue in Section 3.4.3, this is because neither the target profile nor the usage profile remains as a result of the migration. That is to say, this migration has compromised the theory's integrity. Second, I will use the homecoming of evolutionary game theory, which Maynard-Smith criticized, to show in Section 3.4.5 why one should be cautious using a migrated tool. However, I will do so by arguing that Maynard-Smith's critique was not completely justified. These two points are meant to address Statement (***a***) and Statement (***c***), respectively, as I introduced in Section 3.2. Finally, my discussion in Section 3.5 will lead to a refutation of Statement (***b***). That is, contrary to my speculation, in the context of tool migration, losing integrity can, indeed, be optimal and thus desirable.

3.4 The Migration of Game Theory

Game theory was initially formulated to mathematically model strategic interactions between intelligent, rational agents. As defined in game theory, a game is any interaction between two or more people in which each person's payoff is affected by the decisions made by others (Section 3.1.2). Moreover, such a game assumes *complete knowledge* — each player knows the full set of options (or 'moves') for all players, and they all know the payouts to each player in each possible configuration of moves. What they also all know is that all players assume that all players are rational and that all players know that all players assume that all players are rational and so on. That is, there is *common knowledge* of the game and of the rationality of all players.[9] As such, all players will act in the way that takes all other players' potential moves into account in order to maximize their odds of winning. Game theory thus provides tools to mathematically formulate

---

[9] *Common* knowledge is distinguished from *mutual* knowledge. *Mutual* knowledge of P means that all players know that P. *Common* knowledge of P means that not only all players know that P, but they also all know that they all know that P.

strategic interactions and analyze the decisions one can make in those interactions. One such tool is a payoff matrix. For instance, consider an interaction in which one player's gain results in another player's equal loss—called a 'zero-sum game.' The payoff matrix is used to sort out both players' potential moves and the payoffs that follow each of the moves.

Despite being modeled after living-room games to be applied mainly in the study of economic behaviors, game theory has inspired evolutionary biologists to develop an analysis for studying biological evolution. Intriguingly, their framework, which is now known as 'evolutionary game theory' ('EGT'), has subsequently been introduced back to economics and other social sciences for studying aspects of strategic reasoning, including the prisoner's dilemma, among other things. From strategic interactions between rational decision-makers to biological evolution and then back to human interaction, the changes of subject matter are evident. For instance, in social sciences, an act is a result of a player's decision-making. Such a definition simply does not apply in the study of biological evolution. Instead, new definition of the acts, the players, and other theoretical terms were introduced to evolutionary game theory in response to the change of subject matter. Needless to say, when EGT migrates from biology back to the social sciences, significant adjustments to the theory are required, as I will show some of the typical examples in the sections that follow.

3.4.1 The Prisoner's Dilemma

According to Poundstone (1992), the prisoner's dilemma is framed as follows:[10] Two

---

[10] The following account is paraphrased from Poundstone (1992, 118).

members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors admit that they lack sufficient evidence to convict the pair on the principal charge. They plan to get both sentenced to a year in prison on a lesser charge. Simultaneously, the prosecutors offer each prisoner a Faustian bargain. Each prisoner is given the opportunity to either betray the other by testifying that the other committed the crime or to cooperate with the other by remaining silent. The offer is: If prisoner *A* testifies against his partner in crime, he will go free while the partner, prisoner *B*, will get three years in prison on the main charge, and vice versa. If both prisoners testify against each other, both will be sentenced to two years in jail. Both prisoners are given a short time to consider the offer, but in no case may either of them learn what the other has decided until he or she has irrevocably made the decision. Each is informed that the other prisoner is being offered the very same deal, and each of them is concerned only with his or her own welfare—minimizing his or her own prison sentence. The payoff matrix in Table 3.2 summarizes the situation.

Table 3.2: The Payoff Matrix of the Prisoner's Dilemma

	B remains silent, i.e., cooperates	*B* testifies, i.e., defects
A remains silent, i.e, cooperates	1 year (for *A*), 1 year (for *B*)	3 years (for *A*), 0 years (for *B*)
A testifies, i.e., defects	0 years (for *A*), 3 years (for *B*)	2 years (for *A*), 2 years (for *B*)

Prisoner *A* may reason as follows: Suppose I testify and the other prisoner does not. Then I walk away free instead of spending a year in jail. Suppose I testify and the other prisoner does, too. Then I get two years in jail instead of three. Either way, *defect dominates cooperate*. Testifying lessens the sentence regardless of what the other prisoner does. Because defection results in a better payoff than cooperation regardless, which is also true from prisoner *B*'s perspective, it is a dominant move.

Moreover, in this game, neither prisoner could do better by unilaterally switching his or her move from defection. When every player in the game is playing a dominant move, they are said to be in a Nash equilibrium. In a non-cooperative game involving two or more players in which all players have common knowledge of the game and of the rationality of all players, a Nash equilibrium occurs when each player has chosen a move while at the same time no one may benefit by changing his or her moves while other players keep theirs unchanged. In other words, a Nash equilibrium is a set of moves in which every agent's move is optimal. They would choose to maintain their decision given other players doing the same.

This game of prisoners is a dilemma because even though mutual cooperation yields a better result than mutual defection (i.e., 1 year for each against 2 years for each!), from either prisoner's perspective, defecting dominates cooperation. When both prisoners have chosen to cooperate, in hindsight, either of them would have been better off by unilaterally changing his move from cooperation to defection — so as to walk away free instead of serving two years in jail. Thus, mutual cooperation is dominated by mutual defection.

To generalize from the prisoner's situation, any game that has a payoff matrix as seen in Table 3.3 (where *Temptation* > *Reward* > *Punishment* > *Sucker's payoff*) is a variation of the prisoner's dilemma. In particular, the payoff relation $R > P$ implies the analysis that mutual cooperation is superior to mutual defection, whereas the payoff relations $T > R$ and $P > S$ indicate that defection is the dominant move for both players. See Table 3.3 for a summary.

With the prisoner's dilemma as an example, it should be clear that, as a branch of mathematics, game theory's subject matter concerns strategic interactions between rational agents, and the use of the payoff matrix allows for each player to choose an optimal move. Moreover, as one can see from the definition of a Nash equilibrium, a game-theoretic analysis relies on all three assumptions about the players—in other words, complete knowledge, common knowledge, and rationality.[11]

Table 3.3: The General Payoff Matrix of a Prisoner's Dilemma

	Player *B* cooperates	Player *B* defects
Player *A* cooperates	Reward payoff for both (R, R)	Sucker's payoff for *A*; Temptation payoff for *B* (S, T)
Player *A* defects	Temptation payoff for *A*; Sucker's payoff for *B* (S, T)	Punishment payoff for both (P, P)

With similar assumptions, game theory has been used in economics, as well as in other social sciences, to describe, predict, optimize, and explain a variety of human

---

[11] In fact, as a side note, the common knowledge of rationality can be more demanding than necessary when applied to studies of human interactions. For instance, some games are complex enough that it is entirely unreasonable to suppose that all players would have perfect information. Thus, in some branches of game theory, one may study what happens where there are consequent ‘breakdowns’ of rationality—so-called ‘games with bounded rationality.’ In a similar vein, there are games with bounded information (i.e., the players do not necessarily know everything about the structure of the game). Nonetheless, neither games with bounded rationality nor games with bounded information undermine my main point. In all these games, the subject matter remains (i.e., how rational decision-makers strategize in a manner that requires them to make an assessment about what other rational decision-makers will do). It is simply that in some cases, ‘rational’ does not mean ‘fully rational,’ in the sense of ‘always choosing the optimal move.’

interactions in various of domains, such as the economic behaviors of firms, markets, and consumers (e.g., Brandenburger and Nalebuff 1995; Casson 1994) military decisions (Haywood 1954) and international politics (e.g., Snidal 1985).

3.4.2 The Hawk-Dove Game

Game theory was later used in evolutionary biology, when phenotypes or heritable traits were viewed as 'moves,' and individual organisms as embodiments of these moves. In 1972, John Maynard-Smith and George Price borrowed the formalism of a payoff matrix from game theory to mathematically model the evolution of phenotype frequencies in a population of organisms. Terminology that comes out of game theory, such as 'game' or 'payoffs,' can also be seen in this novel use of the theory in biology. Their modeling method assumed that, as the players in a game are in contest with other players, phenotypes are in contest with other phenotypes in a population of organisms. In this sense, a contest is embodied by individual organisms with different phenotypes, and the payoff of a move is formulated using the concept of Darwinian fitness, i.e., the number of copies of itself (offspring) it will leave in the next generation. A game-theoretic analysis in biology typically looks for an 'evolutionarily stable strategy,' which refers to a distribution of phenotypes in a population that is immune to 'invasion' by an initially rare new phenotype. Hence, in other words, the games in biology are no longer strategic interactions between intelligent, rational decision-makers. Instead, these games are but the frequency changes between heritable traits. Similarly, the game-theoretic concept of 'payoffs' has gained a new meaning in biology—Darwinian fitness. Let me illustrate with one of the typical examples: the Hawk-Dove game.

In any population, in contests over resources, hyper-aggressive types of organisms (the Hawks) defeat animals that are peaceful types (the Doves). Thus, one may ask: Why don't we see in nature that the Hawk type organisms take over the population? Maynard Smith and Price (1972) borrowed the payoff-matrix concept from game theory and gave a mathematical account to this question. First, to model this Hawk-Dove interaction as a game, payoffs for a typical example are shown in a matrix in Table 3.4.

Table 3.4: A Payoff Matrix for a Hawk-Dove Game

	Hawk	**Dove**
Hawk	*0*, **0**	*3*, **1**
Dove	*1*, **3**	*2*, **2**

> In this payoff matrix, the payoffs of the moves on the rows are in italics, whereas the payoffs of the moves on the columns are in boldface.

The payoffs assigned to the matrix reflect interactions as follow. Where one is meeting a Hawk, it is better to be a Dove (i.e., *1* > *0* in column 1), and where one is meeting a Dove, it is better to be a Hawk (i.e., *3* > *2* in column 2). Second, a ‘state’ that is resistant to invasion, i.e., an ESS, refers to a distribution of different phenotypes within a population. Such an ESS is a state in which, when embodied by a population in a given environment, none of the individuals in the population may improve their fitness by unilaterally ‘changing their strategy’ (a certain reinterpretation of this phrase is required, which I will address in the next section). It is in this sense that the population is resistant to a few mutants.[12] A population of all Hawks is not stable because no one wins in a

---

[12] Depending on the game, an ESS as a state may consist of players of different ‘types’ (i.e., players who have adopted different strategies, such that none of them can improve

Hawk-Hawk contest (as the payoff profile is *0*, **0**); in such a population, a few Doves, or the 'mutants,' would fare better than the Hawks, or the 'natives.' Through biological reproduction, such a small initial advantage allows for more mutants to enter the population. In a similar way, a population of all Doves would be vulnerable to invasion by a few Hawks. To see how this scenario is mathematically formulated, consider the illustration of an ESS.

The concept of an ESS can be specified formally as follows:[13] Let $\Delta F(s_1, s_2)$ be the change in fitness for an individual following strategy s_1 against an opponent following strategy s_2, both of which are phenotypes, and let $F(s)$ denote the average fitness of an individual following strategy s. Furthermore, suppose that each individual in the population has an initial fitness of F_0. Let σ be an ESS, μ be a 'mutant' strategy en route to invade the population, and p be the proportion of the mutants in the population. Then, the average fitness of a native, $F(\sigma)$, is the sum of its initial fitness and the change in fitness after it interacts with other individuals in the population, including other natives and the mutants (see [3.2]). In a similar manner, the average fitness of a mutant, $F(\mu)$, is the sum of its initial fitness and the change in fitness after it interacts with the natives and other mutants like itself (see [3.3]).

$$F(\sigma) = F_0 + (1-p)\Delta F(\sigma,\sigma) + p\Delta F(\sigma,\mu) \tag{3.2}$$

$$F(\mu) = F_0 + (1-p)\Delta F(\mu,\sigma) + p\Delta F(\mu,\mu) \tag{3.3}$$

---

by unilaterally changing, which is a notion similar to a Nash equilibrium), and at the same time, the population cannot be invaded by new individuals with a new strategy (which is an ESS-like concept).

[13] The following account is paraphrased from Alexander (2009).

For σ to be an ESS, one of two things must be true:

i. The natives fair better when they encounter one another than when the mutants encounter the natives, which can be written as:

$$\Delta F(\sigma,\sigma) > \Delta F(\mu,\sigma) \tag{3.4}$$

ii. The natives do equally well when they encounter one another and when the mutants encounter the natives, but the natives fair better when they encounter the mutants than the mutants do when they encounter one another:

$$\Delta F(\sigma,\sigma) = \Delta F(\mu,\sigma) \text{ and } \Delta F(\sigma,\mu) > \Delta F(\mu,\mu) \tag{3.5}$$

To see how this formula works in the Hawk-Dove game, both the pure hyper-aggressive type of population and the pure peaceful type of population can be invaded by mutants because, contrary to (3.4), $\Delta F(\text{Hawk, Hawk}) < \Delta F(\text{Dove, Hawk})$ and $\Delta F(\text{Dove, Dove}) < \Delta F(\text{Hawk, Dove})$. That is, there is no pure ESS in the Hawk-Dove game, which answers the question why the Hawk type of organisms does not monopolize nature.[14]

3.4.3 An example of tool-transformation

Three major differences stand out between the prisoner's dilemma and the Hawk-Dove game analyses, which, as I argue, jointly qualify the migration of game theory from the social sciences to evolutionary biology a transformative one, i.e., an example of tool-transformation — this conclusion, in turn, supports Statement (***a***) that I laid out in Section 3.2.

First, there is a change of subject matter, i.e., from strategic interaction between decision makers to biological evolution, making the transfer an example of tool

---

[14] This analysis does not rule out the possibility that there may be a hybrid ESS in the Hawk-Dove game.

migration. Second, the target profile of game theory has also changed. Having left the context of the social sciences and entered the context of evolutionary biology, none of the assumptions about the game, the players in a game, and how the payoffs are understood in a game remains the same. Among these, the most important change is the assumption about the 'change of strategy.' Game theory as applied in the social sciences, or classical game theory ('CGT') to be distinguished from EGT, assumes that it is within an individual player's power whether to choose one move over another. Thus, a change of strategy is a result of an individual player's decision-making process.

In contrast, such an assumption (i.e., players being capable of choosing one strategy over another) does not apply to the subject matter of evolutionary biology. As discussed in Section 3.4.2, a 'strategy' in EGT refers to a phenotype that individual organisms display. Moreover, the concept of a payoff, in turn, is re-formulated to be Darwinian Fitness (i.e., the number of copies that it will leave to play in the games of a succeeding generation). Consequently, in EGT, the phrase 'a change of strategy' means a change of the proportion, or frequency, of the phenotypes in a population, which is captured in the mathematical formulations shown in Section 3.4.2. Furthermore, to account for this change in the frequency of phenotypes as a result of the differential fitness, Maynard-Smith assumes that an individual organism passes on its phenotype to its offspring according to the fitness assigned in the payoff matrix. This is known as the heritability assumption about the individuals in EGT. Thus, the eventual change in the frequency of phenotypes is a result of differential biological reproduction and has nothing to do with decision-making.

Third, moreover, a change in the target profile has resulted in a further difference in the use of the two theories, i.e., a change in the usage profile. For instance, the use of EGT no longer carries the ‘optimization use’ as does CGT in the prisoner’s dilemma example. Both CGT and EGT may be used to describe, predict, or even explain certain interactions between two entities (i.e., between rational decision-making agents or between heritable traits).[15] However, in CGT, an analysis of a Nash equilibrium proceeds as if there is complete information about the structure of the game, common knowledge about rationality, and the possession of rationality in individual players.[16] Also, CGT assumes the players as free to making decisions between moves, giving the result of the analysis a prescriptive connotation. To illustrate, recall the prisoner's dilemma. According to CGT, mutual defection dominates mutual cooperation, and if the three assumptions, especially the unbounded rationality assumption, apply, then they *should* both defect. The prisoner’s situation is a dilemma precisely because of the conflict between (1) this prescriptive connotation on the individual player to defect and (2) the fact that mutual cooperation yields a better result for both players.

In contrast, the concept of an ESS—as Maynard-Smith and Price intended—relies

---

[15] It is one thing say that both CGT and EGT may be used to provide explanations regarding their subject matter, it is another to say that they provide the same kind of explanations, which would be a false statement. For instance, the kind of explanations given by CGT would rely on the assumptions about the players’ mental capacity such as reasoning based on complete information, common knowledge, etc., whereas the kind of explanations given by EGT in the context of evolutionary biology would not.

[16] Some might argue that in order for the game-theoretic analysis to be explanatory or predictive, these assumptions do not need to be true of actual individuals. For instance, it could be that human decision-makers are 'hard-wired' to adopt certain Nash equilibrium strategies in the face of strategic interaction not because of any explicit deliberation, in the manner of a game-theorist, but because they are the product of a certain kind of evolution, or simply because over time they have learned that the strategy works. Even in this sense, a Nash equilibrium analysis still depends on the possession of rationality in individual players.

on the heritability of the traits in question and presumes that natural selection through differential reproduction governs the dynamics of the system. In other words, a contest between phenotypes, as seen in a Hawk-Dove game, is not a strategic interaction. Even if there was an ESS in the game, it is not clear in what way that ESS is *prescriptive* as a Nash equilibrium is to a prisoner's dilemma. To put it differently, because individual organisms have no control over the strategy they embody, an ESS is simply a state resistant to mutants as alternative strategies. Thus, unlike a Nash-equilibrium analysis, an ESS analysis may not be used to prescribe courses of action at the level of the individual organisms. This 'prescriptive use' of CGT in the social sciences did not migrate to evolutionary biology. During uprooting, the assumptions regarding the players' ability to make choices between different 'strategies,' are left out of EGT.

In general, in the migration of game theory from social sciences to evolutionary biology, the change of subject matter is evident as it shifted from strategic interactions between rational decision-makers to the dynamics of natural selection. What is also evident is that this change of subject matter is accompanied by a change in the tool's target profile, i.e., the definitions of the theoretical terms and the assumptions of the relations between those terms. Such a change, in turn, affects the tool's usage profile. These definitions and assumptions regarding the decision-making players and their deliberate moves in CGT are left behind, during the uprooting of the tool. In contrast,the heritability assumption regarding phenotypes in EGT introduced as a result of re-situation. The follows that the solution concept in EGT (i.e., the ESS) cannot be used to prescribe an individual's action in the same way as a Nash-equilibrium does in CGT.

Given these changes, which result in the loss of integrity of the tool, it may be a surprising turn of events when EGT is subsequently reintroduced to the social sciences to deal with subject matters again related to rational decision-makers, the homecoming that Maynard-Smith condemned.

3.4.4 Evolutionary Game Theory (EGT): The Prisoner's Dilemma

EGT has found its way back to be used in the social sciences for studying various social phenomena (e.g., Axelrod 1984) or cultural evolution (e.g., Skyrms 2010). As mentioned in Section 3.1.3, Maynard-Smith was doubtful about this homecoming of his work. I recast his qualms using my terminology as follows. By applying EGT to subject matters that are social in nature, social scientists have mischaracterized their phenomenon-of-interest because biological heredity falls short of accounting for human interactions. In this sense, the heritability assumption—one key component of EGT which permits the formulation of differential payoffs between phenotypes in terms of Darwinian fitness—makes EGT a bad tool to apply in the social sciences (see also Grüne-Yanoff 2011b). Let’s take a look at an example of such homecoming: the revisit to the prisoner's dilemma using EGT. This is a story of game theory, having been transformed into EGT in a new scientific context (i.e., biology), returns to its original context. What would the result be when one analyzes the prisoner's dilemma in terms of an evolutionary game? The simple and quick answer is that, just as defection is a Nash equilibrium in the single-shot (i.e., non-evolutionary) game, in an evolutionary version of the game, a population of all defectors is evolutionarily stable.

Taylor and Jonker (1978) tackle the evolutionary version of the prisoner's dilemma as follows:[17] First, assume that the population is large and that individuals encounter one another on a random basis. Simplifying the population in this way, which is common in an ESS analysis, allows one to represent the state of the population by keeping track of the proportions that cooperate and defect, respectively. Let p_c and p_d denote these two proportions. Furthermore, let W_C and W_D denote the average fitness of cooperators and defectors, and let W denote the average fitness of the entire population. Similar to the expressions for fitness in (3.2) and (3.3), W_C, W_D, and W can be expressed in terms of the population proportions and change in fitness:

$$W_C = F_0 + p_c \Delta F(C,C) + p_d \Delta F(C,D) \tag{3.6}$$

$$W_D = F_0 + p_c \Delta F(D,C) + p_d \Delta F(D,D) \tag{3.7}$$

$$W = p_c W_C + p_d W_D \tag{3.8}$$

Second, assume that the proportions of the population between cooperators and defectors in the current generation, p_c and p_d, determine the proportions of the population between cooperators and defectors in the next generation, p'_c and p'_d, based on the average fitness of that proportion over the average fitness of the entire population. This relation can be expressed as:

$$p'_c = \frac{p_c W_C}{W} \qquad \text{and} \qquad p'_d = \frac{p_d W_D}{W}$$

which can be rewritten into the following forms:

$$p'_c - p_c = \frac{p_c (W_C - W)}{W} \qquad \text{and} \tag{3.9}$$

---

[17] The following account is paraphrased from Alexander (2009).

$$p'_d - p_d = \frac{p_d(W_d - W)}{W}\ . \tag{3.10}$$

Using these equations, known as the *replicator dynamics*, one may model the population of cooperators and defectors in the repeated prisoner's dilemma. Recall that in any given prisoner's dilemma, a cooperator receives a *Reward* payoff when encountering another cooperator (i.e., $\Delta F(C, C) = R$) and a *Sucker's* payoffs when encountering a defector (i.e., $\Delta F(C, D) = S$), whereas a defector receives a *Temptation* payoff when encountering a cooperator (i.e., $\Delta F(D, C) = T$) but a *Punishment* payoff when encountering another defector (i.e., $\Delta F(D, D) = P$). Thus, the respective expected fitness of cooperating and defecting is as follows:

$$W_C = F_0 + p_c \Delta F(C, C) + p_d \Delta F(C, D)$$

$$= F_0 + p_c R + p_d S$$

and

$$W_D = F_0 + p_c \Delta F(D, C) + p_d \Delta F(D, D)$$

$$= F_0 + p_c T + p_d P.$$

Recall also that $T > R$ and $P > S$. It follows that $W_D > W_C$ and hence, $W_D > W > W_C$, which entails that

$$\frac{W_D - W}{W} > 0 \qquad \text{and} \qquad \frac{W_C - W}{W} < 0\ .$$

Finally, consider that the proportions of cooperators and defectors in the next generation are given by

$$p'_c = p_c \cdot \frac{(W_C - W)}{W} < 0 \text{ and } \qquad p'_d = p_d \cdot \frac{(W_D - W)}{W} > 0,$$

respectively. Over time, the value of p'_c decreases, whereas the value of p'_d increases, which means that, in a population of both cooperators and defectors, the defectors will eventually take over the entire population. Moreover, this analysis suggests that a population of all cooperators is not resistant against invasion by a few mutants of defectors. Thus, cooperation is not an ESS. In contrast, a few mutants of cooperators may not invade a population of defectors, as the decreasing of the mutants' fitness will drive them into extinction.

The fact that both a Nash equilibrium analysis and an ESS analysis reach the same conclusion about the prisoner's dilemma is not surprising. If a move dominates another at each individual encounter, it will continue to dominate over time. Instead, what makes the homecoming event surprising, germane to my argument, is the change of what counts as ‘payoffs’ in the second time(!) that renders EGT applicable to address the subject matter of CGT.

3.4.5 Misjudgment and mischaracterization

An ESS analysis presupposes that the strategies are inheritable traits (i.e., the heritability assumption). This assumption, as Maynard-Smith points out, makes applying EGT in the social sciences problematic. Maynard-Smith’s criticism turns out not entirely justified. His concern of mischaracterization has been addressed at least by Taylor and Jonker (1978) and Skyrms (2010), respectively. The mathematicians Taylor and Jonker contend that the idea of payoff-as-fitness needs not to be interpreted as resulting exclusively from biological reproduction. As they put it (1978, 146),

> [T]he more fit a strategy is at any moment, the more likely it is to be employed in the future. *The mechanism behind this is either that individuals tend to switch to strategies that are doing well, or that individuals bear offspring who tend to use the same strategies as their parents*, and the fitter the individual, the more numerous his offspring. (Emphasis mine)

Such an increase in the frequency of a particular strategy, they argue, can be understood either as a result of heredity, as Maynard-Smith did, or as a result of social learning (e.g., imitation). Building upon this realization, the 'replicator dynamics' (Section 3.4.4) has been accepted for being an appropriate model both for some biological systems and for learning; the replication of a strategy no longer relies on its previous context—in this case, biological reproduction. Consequently, the modeling of biological reproduction and social imitation with the same mathematical equation (i.e., the replicator dynamics) could be justified. In other words, Taylor and Jonker's effort to re-situate EGT in the social sciences—through giving the formalization of biological reproductive dynamics a new meaning—prevents the users of EGT from mischaracterizing the target phenomenon in their research. Hence, rejecting the tool's novel application in the social sciences on the grounds of the heritability assumption alone would be a misjudgment.

Unfortunately, this is not the end of the story. Owing to Taylor and Jonker's re-interpretation of 'payoffs-as-fitness,' the risk of mischaracterization may *in principle* be set aside. However, well-informed of both Maynard-Smith's worry and the reinterpretation from Taylor and Jonker, Skyrms casts an additional reason for users of EGT in the social sciences to be cautious. According to him (2010, 55):

> The relevant payoffs for cultural evolution may or may not correlate well with

> Darwinian fitness. ... Even if the form of the dynamics is the same for biological and cultural evolution, the substantive conclusions may be different.

That is, each EGT application in the social sciences needs to provide its own interpretation of 'fitness' based on the empirical details of its target phenomenon. In other words, re-situating a research tool does not provide a 'one size fits all' kind of solution to avoid mischaracterization. Skryms concludes, "[c]are in interpretation is required," and it needs to be done empirically based on the context of the application (ibid.). Mischaracterizing a target phenomenon associated with using EGT in the social sciences remains a risk very much *in practice*. Thus, EGT's homecoming makes an example for supporting Statement (***c***), it requires extra care to handle a research tool that has migrated in order to avoid errors such as mischaracterization or misjudgment.

3.4.6 Appropriate Modifications as 'Counter-actors'

One upshot of the above discussion is that the success of applying a migrated tool may not require the integrity of the tool. In other words, losing integrity does not necessarily undermine the usefulness or reliability of a research tool after migration. This observation entails that Statement (***b***) is not the case.

A question naturally arises: If not integrity, what may account for the success of a migrated research tool in the case like game theory? The answer, I argue, lies in the appropriate modification to the tool as a way to *counteract* changes occurring to the tool due to migration. Recall EGT's homecoming that baffled Maynard-Smith. My recount of Taylor and Jonker's episode has shown that applying EGT in the social sciences was not necessarily a 'plug-and-play.' With their re-interpretation of 'fitness,' which constitutes a modification to the tool's target profile, applications of EGT in the social sciences could be plausibly justified. More importantly, this re-interpretation can be viewed as intended to counteract the uprooting of the tool from the context of biology so as to re-situate it to the context of the social sciences. Thus, instead of maintaining integrity, intentionally altering a migrating tool to counteract changes between contexts is key to promising applications of the tool in a new context.

My finding is consistent with the work of some philosophers of science, such as Paul Humphreys and Tarja Knuuttila, who have chosen scientific models and modeling to be the unit of their analysis. Humphreys (2002, 2004) coined the term 'computational templates' for analyzing a relatively small number of mathematical formulas that are used in science to study a variety of phenomena. According to Humphreys, the generality of these templates come partly from their construction and partly from corrections made during each application of them. The former refers to the explicit assumptions that were formulated as part of the template during construction. The latter, in contrast, refers to a "set of not always explicitly formulated suggestions for how to improve" its performance by means of, e.g., idealization, abstraction, approximation (2018, 3). Complementary to Humphreys's analysis, Knuuttila stresses the 'result-driven' approach in the practice of scientific modeling. In such an approach, tweaking an existing template is a necessary strategy when modeling starts from the output and the effects that the models are anticipated to produce instead of otherwise. In her own words (Knuuttila 2011, 268):

> The template that has proven successful in producing certain features of some phenomenon will be applied to other phenomena, often studies within a totally different discipline. [Moreover, i]f a model succeeds in producing the expected

> results or in replicating some features of the phenomenon[,] it provides an interesting starting point for further model building, whose typical aim of which is to correct and adjust the template to better suit the domain it is applied to.

What I conclude from the migration stories of game theory converges nicely with these above observations. For instance, in the homecoming of EGT, seeing the commonality between the replication of a phenotype in the domain of biology and the replication of a strategy in the domain of the social sciences is a stretch of a concept or abstraction. Moreover, such abstraction may be justified by re-interpreting (or, in a sense, tweaking) the meaning of fitness. Thus, losing integrity does not necessarily hinder successfully applying a research tool of an external origin. Quite on the contrary, it is the tweaking of the tool—a process that necessarily undermines the tool's integrity—to properly re-situate it into a new context that seems to predict a promising application. Let us take a look at one more example of game theory's legacy that supports my conclusion.

3.4.7 Evolutionary Game Theory: TIT FOR TAT

The return of EGT to the prisoner's dilemma opens up new possibilities for analysis; in particular, it opens up the possibility of memory-dependent strategies (i.e., 'memory' for the players). Recall the prisoner's dilemma. It is a dilemma because it "embodies the tension between individual rationality (reflected in the incentive of both sides to be selfish) and group rationality (reflected in the higher payoff to both sides for mutual cooperation over mutual defection)" (Axelrod 1980, 4). But as Robert Axelrod (1980) points out, this analysis rests primarily on either the 'one-off' interaction between the players (as in CGT) or the 'random pairing' between the players (as in EGT). Modifying the assumption of how players interact may consequently change the result of the analysis. After all, Axelrod argues, human strategic interactions are neither isolated nor random, and thus "making effective choices … requires insight into the structural implications [i.e., the history of the game thus far] of strategic interaction" (ibid., 4). That is, players who are capable of identifying their opponent and remembering the history with this opponent in the previous rounds of the game will make their strategic move according to that history.

In addition to 'defect' and 'cooperate,' many other strategies are introduced to the analysis due to this new assumption. One such strategy is called 'TIT FOR TAT.' As a decision rule for making a move in each round of the game, TIT FOR TAT always starts out by cooperating, and in all the following matches, it simply *echoes* what its opponent did in the immediately previous match.

It turns out that TIT FOR TAT is an ESS that can be applied to the memory-dependent version of the prisoner's dilemma against the opponent who always defects ('ALL D') (Axelrod 1984). A population of individuals playing TIT FOR TAT cannot be invaded by a few mutants that play ALL D. Individuals playing ALL D will always receive punishment payoff, which is lower than the reward payoff that individuals playing TIT FOR TAT receive. For this reason, a few mutants of ALL D players will drive themselves into extinction. However, as one may wonder, can the population of ALL D players be invaded by players who follow TIT FOR TAT? The answer depends on two factors: the initial size of the mutants and the pairing of mutants with other players. A population of ALL D players is resistant to mutants of TIT FOR TAT players when the number of mutants is small and the pairing is random. TIT FOR TAT always starts out by cooperating, and so it can be exploited by ALL D. However, if there are enough mutants of TIT FOR TAT that tend to be paired together in an ALL D native population, it is possible that over generations, TIT FOR TAT will take over the entire population.

Notice that this idea of a ‘memory-dependent strategy’ would not make sense in a purely biological context, because evolution is not an explicitly *strategic* contest. But once EGT has made its way back into the social sciences, then there is room for such memory-dependent, strategic thinking. Hence, what we see in an ESS analysis of TIT FOR TAT is not only a case where a tool ‘left to return home’ but also a case where the tool was modified when it left (adding the concept of evolution while leaving out ‘strategic thinking’) and then modified *again* upon its return (introducing memory-dependent strategic thinking). In other words, what prima facie seems to be the borrowing of one research tool across multiple disciplines turns out to be a natural history of the tool’s development and evolution, eventually leading to several different target profiles of the research tool. Such a discovery, again, supports that losing integrity can contribute to the usefulness of a research tool after migration. It is the modification of the tool that seems to play a pivotal role.

3.5 Tool Migration Risks

Studying the cross-disciplinary use of formal constructs in science as tool migration has the potential to reveal the risks related to using a migrated tool and how scientists deal with these risks. Generalizing from the example of EGT, one notes that a migrated tool may lose a particular usage (i.e., the prescriptive interpretation of the solutions) due to a change of its target profile. Moreover, to counteract a change of the subject matter, scientists can modify a migrated tool so that it may be used appropriately in a new context, e.g., free from mischaracterizing a new target phenomenon. Sometimes, the modification needs to be performed according to the context of each application. Arguably, such a counter-acting modification requires the user to *be aware of the need* to modify the migrating tool. One might suggest that a migrated tool could be 'unintentionally' modified and working successfully in a new context. While I do not rule out the possibility of an unintended positive consequence of tool migration, such a success is out of luck, and when the user is indeed ignorant as to why the migrated tool works in his or her epistemic pursuit, it is doubtful whether the research indeed results in genuine knowledge.

Moreover, and importantly, applying a migrated tool without being aware of the need to counteract puts the user at the risk of error. For instance, recall the migration of game theory from the social sciences to biology. By taking away the assumption of 'choice' from the relation between a player and a move, an ESS may no longer have a prescriptive interpretation as does a Nash equilibrium. A prescriptive interpretation of an ESS could thus overstep the bounds of the tool, constituting a misinterpretation of the result.

Finally, the purpose of pointing out these risks of mischaracterization and misinterpretation associated with tool migration is not to argue against using research tools that have migrated. As I commented on Maynard-Smith's concern about EGT's homecoming, hast preclusion of using such tools could lead to misjudgment, in the sense of falsely rejecting a viable research tool. Instead, what I wish to do is to motivate a study of tool migration. There remain questions to be asked about tool migration as a vibrant feature of scientific practice. For instance, what does it look like in a case where a tool migrates across disciplines while its target profile remains critically similar? Can the tool's usage profile change without critically changing its target profile? In the next two chapters, I will address these questions by going over the case of formal language theory.

Chapter 4

Formal Language Theory — From Linguistics to Computer Science

Formal language theory ('FLT' henceforth), including its components such as automata theory and the Chomsky hierarchy, is the study of mathematically defined languages. Initially formulated by Chomsky in the 1950s to investigate syntactic regularities of natural languages, FLT remains a branch of mathematics and linguistics (Levelt 2008). However, much of its subsequent development in the 1960s and 1970s was achieved in theoretical computer science, making its presence more dominant in computer science than in linguistics (Greibach 1981). Since 2004, the topic has seen a surge of research, in which psychologists apply FLT to study nonhuman animals' abilities to learn artificial grammars — the objective is to understand both the evolution of, and neural substrate for, the faculty of human language.[18] I contrast the cross-disciplinary applications of FLT in the next two chapters. In this chapter, I discuss the use of FLT in linguistics and computer science, respectively. In Chapter 5, I discuss its more recent applications in the study of the evolution of the mind.

I begin this chapter with a section of preliminary discussion about the theory of

---

[18] For instance, as Fitch (2014) reports, Abe et al 2011; Bahlmann et al 2008; de Vries et al 2008; Fitch and Hauser 2004; Gentner et al 2006; Hochmann et al 2008; Marcus 2006; Perruchet and Rey 2005; Rey et al 2012; Stobbe et al 2012; Udden et al 2012; van Heijiningen et al 2009.

formal languages. In Section 4.1, I briefly review some key components of FLT relevant to its application in cognitive biology, such as the formal systems of grammars and abstract machines and a classification scheme called the Chomsky hierarchy. In Section 4.2, I focus on FLT's initial applications in linguistics as presented by Chomsky. His application includes using FLT to explain semantic ambiguity, to formalize linguistic analysis, and to narrow down to a method for modeling the syntax of natural languages. Moreover, some of the concepts that I introduce in Section 4.2 are crucial to arguments that I go through in later sections, as well as in the next chapter. I will refer back to these subsections as needed. In Section 4.3, I describe early applications of FLT in computer science, such as the design of programs, programming languages, and compilers. In Section 4.4, based on the framework of tool migration that I developed in Chapter 3, I analyze these two cases, i.e., the initial, established use of FLT in linguistics and the then-novel use of FLT in computer science. In particular, I discuss in what sense (1) FLT can be viewed as a research tool and (2) FLT has migrated from linguistics to computer science. I then argue that (3) this migration is a case of tool-adaptation (i.e., the migration changes FLT's usage profile without substantially changing its target profile).

4.1 Theory of Formal Languages

4.1.1 An Overview

The building blocks of FLT include the mathematical definitions of an 'alphabet,' a 'sentence,' a 'language,' a 'grammar,' and an 'automaton,' all of which except for the automaton were originated in Chomsky's (1956, 1959b) study of natural languages (e.g., Griebach 1981, Hopcroft 1987, Levelt 2008). In FLT, an alphabet (also called a ‘vocabulary’) is any finite set of symbols, a sentence is any finite string composed of the elements from a given alphabet according to the grammar of a language, and a language is any set of such sentences. Most important, a grammar is a formal system which, like other formal systems such as propositional logic, transforms “a certain input into a particular output by means of completely explicit, mechanically applicable rules” (Levelt 2008, 2). In this context, input and output are themselves also strings of symbols from the alphabet. In other words, a grammar is a formal device that takes as input the elements of an alphabet and produces as output the sentences of a language.

To put these above descriptions in formal terms, let L be a language, V be an alphabet or a set of vocabulary, and V^* be a set whose members are strings composed by concatenating zero or more symbols in V. Then $L \subseteq V^*$. Moverover, let ε be the string of zero elements, let V^+ be V^* without ε, and P be a set of rules. Then the cartesian product $V^+ \times V^*$ is a set of all possible ordered pairs of strings over V, and consequently, $P \subset V^+ \times V^*$ shows how a string is rewritten as another string by a single application of a rule in P.

Similar to grammars, automata are also formal systems. While a grammar is constructed to produce a formal language, an automaton is constructed to determine whether a given string belongs to a particular formal language. An automaton takes as input a string of symbols and gives a ‘Yes’ or ‘No’ verdict regarding the membership of the string to a particular language. In other words, grammars are the ‘generators’ of formal languages, whereas automata are the ‘recognizers’ of formal languages. For this reason, both grammars and automata are legitimate means to describe a language.

There are other ways to describe a formal language. For instance, one may list all sentences of the language; this method is apparently limited to languages with finite (and small enough) number of sentences. In contrast, for languages with an infinite number of sentences, one may write out sufficient sentences for the pattern of the sentences to be made clear, then use '...' to indicate that the pattern will continue indefinitely, such as {*ab*, *aabb*, *aaabbb*, ...}. One may also use set-theoretic notation such as $\{a^n b^n : n \geq 1\}$ to refer to this aforementioned set, whose members are made of strings of a's and b's where the number of a's is equal to the number of b's.

One remark before I introduce the formal systems of a grammar and automaton: In an important sense, both grammars and automata are 'language-specific.' The grammar of a language is assumed to generate *all and only* sentences of that language (nothing else!). An automaton of a language recognizes also *all and only* sentences of a particular language (nothing else). For this reason, to describe a language, one can either specify the grammar of that language or supply an automaton for the said language. Both are acceptable means of describing a formal language. The interchangeability between grammars and automata has led some psycholinguists to suggest that "the distinction between generating [a language] and accepting [it]" is primarily conceptual, and "less fundamental than it may at first appear" (Levelt 2008, 2). That is, rather than using a formal grammar, one can "very well use an automaton ... as a model for a theory of natural language"; although this has in fact been done, "the generative grammar remains the preferred model" (ibid.). According to Levelt, the dichotomy between grammars and automata "has a natural interpretation with reference to speaker-hearer models" (ibid.).[19]

The study of automata (e.g., Turing 1936) predates Chomsky's study of formal

---

[19] As I discuss in Section 4.2.3, Chomsky (1959b) does not endorse such an interpretation.

grammars, but soon after the Chomsky hierarchy was introduced (1959b), scholars began to see the connection between the two. What initially appeared to be two independent areas of study conjoined under the name of 'formal language theory.' In what follows, I introduce two main elements of FLT related to our present focus: the 'phrase structure grammars' (for specifying formal languages) and the Chomsky hierarchy (for classifying formal languages based on their expressive power).

4.1.2 Phrase Structure Grammars

Constructed by Chomsky to study natural languages, a phrase structure grammar ('PSG') specifies a formal language by defining an alphabet and a set of rules for rewriting strings composed by symbols from the alphabet.

The alphabet of PSG, denoted by V, distinguishes between the 'terminal' vocabulary, V_T, and the 'non-terminal' vocabulary, V_N. V_T, is the set of terminal elements, i.e., words, with which the sentences of a language may be constructed. Elements of V_T are usually denoted by lower case letters (i.e., *a*, *b*, *c*…). V_N consists of non-terminal elements, which are categorical symbols used either in the beginning of the sentence production or during the process of the sentences production; they never appear in the sentences of the language. Elements of V_N are typically indicated by upper case Latin letters (i.e., *S*, *A*, *B*, …). For example, the symbol *S* is a special, non-terminal symbol; it is usually called the 'start' or the 'sentence' symbol. Note that a string can be composed of zero or more symbols from either V_N or V_T or both, but V_N and V_T do not have any elements in common, i.e., $V_N \cap V_T = \{\}$. The rules of PSG for rewriting strings, also called the 'rewrite rules' or 'productions,' are in the form of $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ where both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

stand for any sequence of (terminal or non-terminal) symbols and $\boldsymbol{\alpha} \in V^+$ and $\boldsymbol{\beta} \in V^*$. ‘ → ’ is called the production arrow.

With an alphabet and a set of rules, a PSG can be used to specify in a stepwise manner how any given sentence of language may be *derived*. For example, let a grammar $G = (V_N, V_T, P, S)$, where $V_N = \{S\}$, $V_T = \{a, b\}$, $P = \{S \rightarrow aSb, S \rightarrow ab\}$.[20] G generates the language $L(G) = \{ab, aabb, aaabbb, ...\}$, and the two productions shown in (4.1) are said to be able to derive every sentence in $L(G)$. To illustrate, (4.2) shows a derivation of the string *aaaabbbb*.

(4.1)

i. $S \rightarrow aSb$

ii. $S \rightarrow ab$

(4.2)

$\#^\frown S^\frown\#$	Start
$\#^\frown S^\frown\# \Rightarrow \#^\frown \boldsymbol{aSb}^\frown\#$	Applying rule $S \rightarrow aSb$
$\#^\frown aSb^\frown\# \Rightarrow \#^\frown a^\frown \boldsymbol{aSb}^\frown b^\frown\#$	Applying rule $S \rightarrow aSb$
$\#^\frown aaSbb^\frown\# \Rightarrow \#^\frown aa^\frown \boldsymbol{aSb}^\frown bb^\frown\#$	Applying rule $S \rightarrow aSb$
$\#^\frown aaaSbbb^\frown\# \Rightarrow \#^\frown aaa^\frown \boldsymbol{ab}^\frown bbb^\frown\#$	Applying rule $S \rightarrow ab$

In (4.2), the column on the left indicates the steps and the strings produced in each of the steps. The column on the right indicates the rules applied in those corresponding steps. ‘#’ stands for the beginning or the end of a string, and ‘ $^\frown$ ’ stands for concatenation.[21] A

---

[20] Fitch and Hauser (2004) later call this grammar ‘the A^nB^n grammar’ in their experiment on artificial grammar learning — an episode which I feature in Chapters 5 and 6.

[21] These symbols were introduced in Chomsky (1956a) but omitted by most of the later authors in their introduction to the theory of formal languages.

derivation starts with an S, the 'sentence' symbol, and when it arrives at a sequence of non-terminal symbols, it comes to a stop. In particular, one applies the rule in (4.1i) to replace S with ***aSb***, where ⇒ stands for 'replace,' and the symbols shown in boldface in each step indicate the sequence generated in that step.[22] The whole process as seen in (4.2) is conventionally shown in a more concise way:[23] $\{S \Rightarrow aSb \Rightarrow aaSbba \Rightarrow aaSbbb \Rightarrow aaaSbbb \Rightarrow aaaabbbb\}$, which is called a 'derivation of *aaaabbbb*.'

In FLT, a derivation can be represented in a tree diagram. See Fig 4.1 for an example.[24] In a tree diagram, each of the non-terminals is represented as a node that extends further down until the branch reaches to a terminal (or leaf) node. The resulting string of a derivation is taken by reading the leaf nodes in left-to-right order. As I discuss in Chapter 6, this tree-shape representation of the derivation of sentences plays a crucial role in some cognitive biologists' theorization of mind and brain (e.g., Fitch 2014).

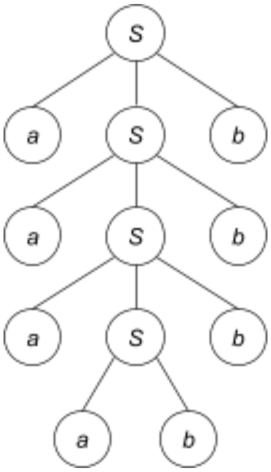


Fig. 4.1 A Derivation Tree

---

[22] Because the rewriting of S into ***aSb*** occurs at the center of the string, in linguistics, this process is also referred to as 'center-embedding.'

[23] C.f. Parkes (2002).

[24] C.f. Parkes (2002, 38).

To summarize, in the theory of formal languages, a PSG $G = (V_N, V_T, P, S)$ is a formal system consisting of a non-terminal vocabulary V_N, terminal vocabulary V_T, a set of productions P, and a special start symbol S. Productions are rewrite rules which specify which strings in V^+ can be rewritten to strings in V^* stepwise. Note that the productions in a PSG are in the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ where $\boldsymbol{\alpha} \in V^+$ and $\boldsymbol{\beta} \in V^*$ (Levelt 2008, 4). Thus, for all PSG productions in the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$, $\boldsymbol{\alpha} \in (V_N \cup V_T)^+$ and $\boldsymbol{\beta} \in (V_N \cup V_T)^*$. That is, while the empty string ε can be on the right-hand side of a production, it cannot be on the left-hand side of a production. This form of PSG productions will be the basis of the Chomsky hierarchy that I discuss in the section that follow.

4.1.3 The Chomsky Hierarchy of Formal Grammars and Languages

In its initial form, the Chomsky hierarchy (1959b) is a classification of PSGs and the formal languages these grammars generate, now called 'context-sensitive,' 'context-free,' and 'regular.' Although the construct of PSG was first published in 1956, it wasn't until 1959 that Chomsky (1959b) constructed this hierarchy now bearing his name. It "placed formal languages in a mathematically attractive light and began to draw theoretically oriented adherents" (Ginsburg 1980, 4). The way Chomsky classifies different classes of grammars is based on three increasingly restrictive conditions on the productions, as follows:[25]

First limiting condition: For every production $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ in P, $|\boldsymbol{\alpha}| \leq |\boldsymbol{\beta}|$, i.e., the number of symbols on the left-hand side of the production is smaller than or equal to the number of

---

[25] This review is based on Levelt (2008, 10); c.f., Partee, Meulen, and Wall (1990) for a slightly different review of the Chomsky hierarchy.

the symbols on the right-hand side of the production. That is, the resulting string $\boldsymbol{\beta}$ will always be longer or equal to the length of the string $\boldsymbol{\alpha}$ before the rewrite. Thus, applying the productions of this grammar would not result in a decrease of the string length. Note that the symbol of empty string ε has length = 1. Hence, $S \rightarrow \varepsilon$ satisfies this condition.

Second limiting condition: For every production $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ in P, (1) $\boldsymbol{\alpha}$ consists of only one non-terminal symbol, i.e., $\boldsymbol{\alpha} \in V_N$ and $|\boldsymbol{\alpha}| = 1$, and (2) $\boldsymbol{\beta} \neq \varepsilon$, i.e., $\boldsymbol{\beta} \in V^+$.

Third limiting condition: For every production $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ in P, (1) $|\boldsymbol{\alpha}| = 1$, $\boldsymbol{\alpha} \in V_{N,}$ and (2) $\boldsymbol{\beta}$ has the form $\boldsymbol{\gamma}$ or $\boldsymbol{\gamma\Psi}$, where $\boldsymbol{\gamma} \in V_T$, $\boldsymbol{\Psi} \in V_N$, and $|\boldsymbol{\gamma}| = |\boldsymbol{\Psi}| = 1$, e.g., $A \rightarrow a$, $A \rightarrow aB$.

With these limiting conditions, four types of PSGs may be classified as follows:

Type 0 grammars (also called unrestricted) are grammars to which none of the three limiting conditions applies. In other words, all of the productions are of the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$, where $\boldsymbol{\alpha} \in (V_N \cup V_T)^+$ and $\boldsymbol{\beta} \in (V_N \cup V_T)^*$. Except for ε, which cannot appear on the left-hand side, all symbols (terminal or nonterminal) are allowed to be on either side of the production arrow.[26] For instance, $aXYpq \rightarrow aZq$ conforms to this pattern.

Type 1 grammars (also called the context-sensitive) are grammars to which the first limiting condition applied. All the productions of Type 0 with the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ where $|\boldsymbol{\alpha}| \leq |\boldsymbol{\beta}|$ belong to Type 1. That is, Type 1 grammars constitute a strict subset of the Type 0 grammars. The Type 0 grammar that are not Type 1 are, for example, those with at least

---

[26] Some additional conditions are required for a Type 0 grammar to be able to actually generate sentences, such as (1) having to include a rule in which $\boldsymbol{\alpha}$ is the start symbol and (2) at least one symbol in $\boldsymbol{\beta}$ needs to be a non-terminal.

one production where $|\boldsymbol{\alpha}| > |\boldsymbol{\beta}|$.

Type 2 grammars (also called the context-free) are grammars restricted by the second limiting condition. All productions of Type 1 grammars with the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ where $|\boldsymbol{\alpha}| = 1$, $\boldsymbol{\alpha} \in V_N$, and $|\boldsymbol{\beta}| \geq 1$ belong to Type 2. Note that the second limiting condition entails the first, i.e., from $|\boldsymbol{\alpha}| = 1$ and $|\boldsymbol{\beta}| \geq 1$, it follows that $|\boldsymbol{\alpha}| \leq |\boldsymbol{\beta}|$. Thus, context-free grammars are also context-sensitive grammars. But the inverse is not true; the Type 1 grammars with at least one production where $\boldsymbol{\beta} = \varepsilon$ (e.g., $S \rightarrow \varepsilon$) would be excluded from also being Type 2. Thus, the Type 2 grammars constitute a strict subset of the Type 1 grammars.

Type 3 grammars (also called regular grammars) are grammars restricted by the third limiting condition. The productions of this type are in the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ where $\boldsymbol{\alpha} \in V_N$ and $\boldsymbol{\beta}$ has the form of either $\boldsymbol{\gamma}$ or $\boldsymbol{\gamma\Psi}$ where $\boldsymbol{\gamma} \in V_T$, $\boldsymbol{\Psi} \in V_N$, and $|\boldsymbol{\alpha}| = |\boldsymbol{\gamma}| = |\boldsymbol{\Psi}| = 1$, e.g., $A \rightarrow a$, $A \rightarrow aB$. Note that the third limiting condition entails the second. That is, from $|\boldsymbol{\alpha}| = |\boldsymbol{\gamma}| = |\boldsymbol{\Psi}| = 1$ and $|\boldsymbol{\gamma}| + |\boldsymbol{\Psi}| = |\boldsymbol{\beta}|$, it follows that $|\boldsymbol{\alpha}| = 1$ and $|\boldsymbol{\beta}| \geq 1$. That is, regular grammars are also context-free, but the converse is not true. There are context-free grammars that are not regular, e.g., those with at least one production where $|\boldsymbol{\beta}| > 2$ (e.g., $S \rightarrow Sab$, $S \rightarrow aSb$. Note, we will encounter these two production rules later in Chapter 6). Thus, the Type 3 grammars form a strict subset of the Type 2 grammars.

These four types of grammars form a containment hierarchy which applies not only to the formal grammars but also to the formal languages they generate. First, in terms of the grammars, the four types of grammars are constructed based solely on the increasingly strict requirements on the forms of the productions. Thus, any grammar of a more restrictive type will satisfy the restrictions of any other types that are less restrictive but not vice versa. The resulting scheme is a containment hierarchy of formal grammars. That being said, conventionally, once a grammar (which may contain more than one production) satisfies the restrictions of Type 3, it is instantly classified as a Type 3 grammar.[27] Only when it does not (i.e., at least one production failing to meet the third limiting condition), will it be checked as to whether it qualifies as a Type 2 grammar. If it does, then it is classified as so, and this pattern repeats until possibly reaching Type 0, the unrestricted grammars. That is, by definition, any phrase structure grammar fulfills the requirements of being in the class of Type 0 grammars, but by convention, grammars are classified based on the extra restrictions they also display in the patterns of their productions. In computer science, such a convention is of practical importance, e.g., to the design of compilers. In Section 4.3, I discuss the use of the Chomsky hierarchy in computer science.

Second, in terms of languages, the formal languages generated by each of the four classes of grammars also form a containment hierarchy. In its standard formulation (as reviewed in Greibach 1981), the four classes of languages in the Chomsky hierarchy are: recursively enumerable (i.e., computable), context-sensitive, context-free, and regular, respectively. Among these four classes of languages, both context-sensitive and context-free were first described by Chomsky (1956, 1959b) as produced by the Type 1 and Type 2 grammars. In contrast, recursively enumerable languages were first described by Emil Post (1943, 1947) using the method of 'rewriting systems.' Post's rewriting system became the basis of Chomsky's PSGs, i.e., the Type 0 grammars, making

---

[27] Parkes 2002, 31.

recursively enumerable languages the outermost member in the hierarchy (Greibach 1981). Finally, regular languages were defined by Stephen Cole Kleene (1951, 1956) using the formulation based on McCulloch and Pitts' (1943) mathematical model of nervous activity. Chomsky and Miller (1958) showed that the Type 3 grammars can produce regular languages. See Fig 4.2 for a graphical representation of the containment hierarchy of four formal languages.

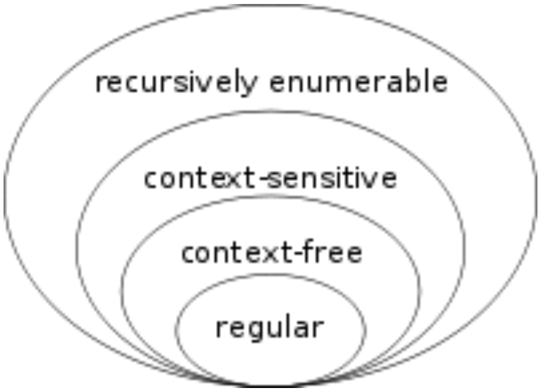


Fig 4.2 The Hierarchy of Formal Languages

4.1.4 Automata and Their Corresponding Formal Languages and Grammars

Although FLT is said to have originated in Chomsky's work in the 1950s, some of its components had been circulating before then, such as the automata theory. Automata are mathematical models of abstract computers that "transform information from one form into another on the basis of predetermined instructions" (Encyclopaedia Britannica).[28] The first of such models is commonly taken to be the Turing Machine (Turing 1936).[29]

A Turing machine consists of a 'head' and a 'tape' divided into cells that extend

---

[28] https://www.britannica.com/technology/automaton Access Date: February 19, 2019.

[29] Other candidates include Church's lambda calculus (1936).

infinitely in both directions. It also has an internal state (one member of some set of possible states) and a program, i.e., a set of instructions. Each cell contains a single symbol from some alphabet of symbols. The head can read the symbol, and then, according to the program, perform any combination (including none) of acts as follow: Based on the information that is written in the cell and the internal state of the machine, a Turing machine will (1) write a new symbol on the cell, (2) move one cell left or right on the tape, or (3) change its internal state. There is a special internal state, the 'halting state,' and once the machine reaches this state (if it ever does), it performs no further actions. In a sense, a Turing machine can be thought of as a 'state machine' with a working tape of infinite length.

A working tape of an infinite length is of theoretical importance. With the formulation of an abstract machine that now bears his name, Turing was aiming to address the question about the limitation of compatibility, i.e., whether there exist uncomputable mathematical problems, also known as David Hilbert's *Entscheidungsproblem* or the decision problem. An infinite working tape allows Turing to explore the limitation of computability without being confined by other limitations such as memory (or time).

Other subsequently constructed automata are also state machines: They differ in the presence of an explicit storage and the capacity of such a storage. For instance, finite-state automata are state machines without a working tape of any sort. In contrast, the pushdown automata and the linear-bounded automata are (just like the Turing machines) essentially a finite-state automata equipped with some explicit storage.

The study of finite-state automata is thought to have originated in McCulloch and

Pitts' (1943) mathematical models for neural networks and later developed by Kleene (1952, 1956) (Levelt 2008). A finite-state automaton is defined by a list of finite states, including an initial state, and the conditions for transitioning between the states. The stark contrast between a finite-state automaton and other types of automata is its lack of explicit storage of information. As Kleene defines (1956, 3):

> A nerve net is an arrangement of a finite number of neurons in which each endbulb of any neuron is adjacent to the soma of not more than one neuron (the same or another) the separating gap is a synapse. Each endbulb is either excitatory or inhibitory (not both).

Kleene calls the neurons without endbulbs 'input neurons,' and the others, 'inner neurons.' Each neuron is assumed to be either firing or not firing. For an input neuron, the behavior (firing or not) at any given time is determined by conditions outside the net. "One can suppose each is impinged on by a sensory receptor organ, which under suitable conditions in the environment causes the neuron to fire" (Kleene 1956, 3-4). Following the McCulloch-Pitts model, Kleene's automata naturally do not come with explicit storage for information, i.e., memory. His investigation of the McCulloch-Pitts nerve nets was "only partly for their own sake as providing a simplified model of nervous activity, but also as an illustration of the general theory of automata, including robots, computing machines and the like. Nonetheless, "to prevent misunderstanding," he writes "memory can be explained on the basis of reverberating cycles of nerve impulses. This seems a plausible explanation for short-term memories" (ibid., 2).

The presence of an explicit storage distinguishes finite-state automata from all other automata. In pushdown automata, the storage is called a 'last-in-first-out' stack. It is a highly restricted form of storage: At any given time, the one and only one item in the storage that can be retrieved is the most recent item that was added to the storage. To illustrate, consider several cups stacked up sitting on a table. Consider that you have only one hand available (because your other hand is busy, say, holding heavy books), and you are allowed to take only one cup at a time. Thus, at any given time, only the cup on top of the stack is retrievable to you. Similarly, at any given time, to add a new cup to the stack, you can do so only by placing it on top of the stack. Consequently, the earlier a cup is added to the stack, the lower its position in the stack (i.e., closer to the table); the lower its position, the more cups one needs to remove before retrieving it. In the linear-bounded automata, the storage is a working tape that operates just like the one in a Turing machine except its length is limited (i.e., its length is predetermined by a constant times the length of the input).

Together, Turing machines, linear-bounded automata, pushdown automata, finite-state automata form a hierarchy of abstract machines based on the increasingly heavy restrictions on the storage. As said earlier, a Turing machine has an unlimited length for storage, making it the least restricted type of automata. An linear-bounded automaton is less restricted than a pushdown automaton because, everything else being equal, on a working tape, any item can be retrieved at any time, but on a stack, only the last stored item can be retrieved at a given time. Finally, a pushdown automaton is less restricted than a finite-state automaton because, everything else being equal, the former has an explicit storage for information, whereas the latter does not.

Each type of automata is found ‘equivalent’ to one of the four classes of formal grammars on the Chomsky hierarchy. A grammar and an automaton are ‘equivalent’ just in case the languages produced by the grammar are the same languages that can be recognized by the automaton. In what follows, I list the equivalent pairs of grammars (the languages they produce) and automata. In a sense, the Chomsky hierarchy provided a scheme for conceptualizing different formal systems (e.g., mathematical models of computation, neural nets, grammars) that might not initially be perceived to be related to one another.

Type 0 unrestricted grammars (recursively enumerable languages) and Turing machines: Post (1947) showed his rewriting systems can produce those languages that are recognizable by Turing machines, establishing the outermost class of the hierarchy (Greibach 1981). It has been shown (Davis 1958, Chomsky 1959b, 1963) that Type 0 grammars produce recursively enumerable languages.[30] Thus, Turing machines are equivalent to Type 0 grammars: Type 0 unrestricted grammars are the generator for, and the Turing machines are the recognizer of, the recursively enumerable languages. Basically, Type 0 grammars include all formal grammars; they generate all the formal languages that could be computed at all.

Type 1 context-sensitive grammars (context-sensitive languages) and the linear-bounded automata: Chomsky (1959b) defined Type 1 grammars by placing restrictions to phrase structure grammars (which was based on Post's rewrite systems). The equivalence between a context-sensitive grammar and the linear-bounded automata was proved in the 1960s by Landweber (1963) and Kuroda (1964) (Levelt 2008).

---

[30] A summary of the proofs can be found in Levelt 2008, 100-107.

Type 2 context-free grammars (context-free languages) and the pushdown automata: According to Levelt (2008), the notion of a 'pushdown store' was introduced by Newell, Shaw, and Simon (1959), but it was Oettinger (1961) who first drew the relation between pushdown automata and formal languages. Chomsky (1963) and Evey (1963) "more or less independently" formulated the relation between context-free grammars and pushdown automata (Levelt 2008, 126).

Type 3 regular grammars (regular languages) and the finite-state automata: According to Levelt (2008), Chomsky and Miller (1958) showed the equivalence of finite-state automata and regular grammars.

See Fig. 4.3 for a summary.

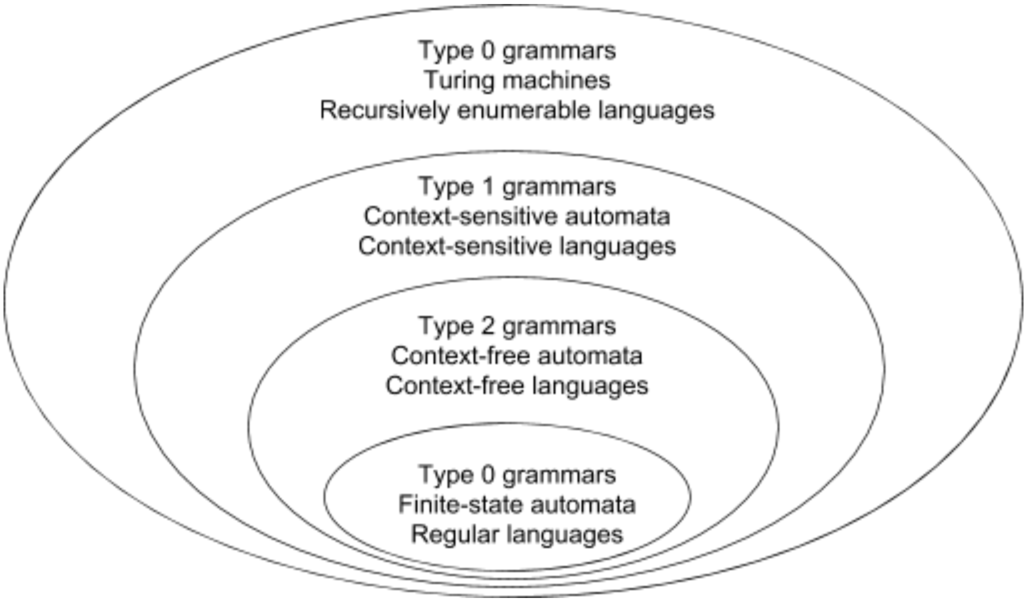


Fig. 4.3 The Chomsky Hierarchy of Formal Grammars, Automata, and Languages

4.2 Theory of Formal Languages in Linguistics

In this section, I discuss three examples of how FLT has been used in linguistics. In

Section 4.2.1, as the first example, I discuss how Chomsky explains semantic ambiguity using the tree-like diagrams to represent sentence derivation. This application of FLT shows Chomsky's attempts to justify using PSG to study natural languages. In Section 4.2.2, as the second example, I review Greibach's (1981) account of PSG as Chomsky's attempt to formalize research methods in linguistics called 'immediate constituent analysis.' In Section 4.2.3, I discuss the 'correspondence hypothesis' related to the analysis in linguistics and psycholinguistics. Some linguists thought that their analysis of the construction or parsing of sentences, to an extent, informs us the speaker-hearer's psychological processes of utterance production or comprehension. In contrast, Chomsky did not endorse such an interpretation. In Section 4.2.4, as the third example, I discuss what Chomsky actually did with the Chomsky hierarchy: he used it to narrow down a plausible model of the syntax of natural languages. In particular, he argues that English cannot be modeled using a finite-state grammar, i.e., English is 'supra regular.' In Section 4.2.5, I conclude this section by contrasting Chomsky's conclusion that the syntax of natural languages is supra regular and the hypothesis that the faculty of human language is supra regular, which I call the 'supra regular hypothesis.'

Chomsky's goal to model natural languages never delivered, though it is worth noting that he (1959b) was not fully endorsing the correspondence hypothesis. Nonetheless, his conclusion that English is supra-regular has been interpreted as a hypothesis concerning both psycholinguistics and neurolinguistics and recently put to test by cognitive biologists (Fitch and Hauser 2004, Fitch 2014).

4.2.1 Explaining Semantic Ambiguity

Chomsky shows that PSG and the tree diagrams can explain why some sentences in English are ambiguous. Consider the sentence ‘they are flying planes.’ One could take the third-person plural pronoun, ‘they,’ as referring to the same objects referred to by the word ‘planes’ in the same sentence. Alternatively, one could understand ‘they’ as referring to some pilots who are flying planes. Because linguistic ambiguity like this exists in English, Chomsky (1956, 118) argues that “the grammar of English will certainly have to contain such rules” as I show in (4.3).

(4.3)

Sentence → Noun Phrase⁀Verb Phrase

Verb Phrase → Verb⁀Noun Phrase

Verb → are⁀flying

Verb → are

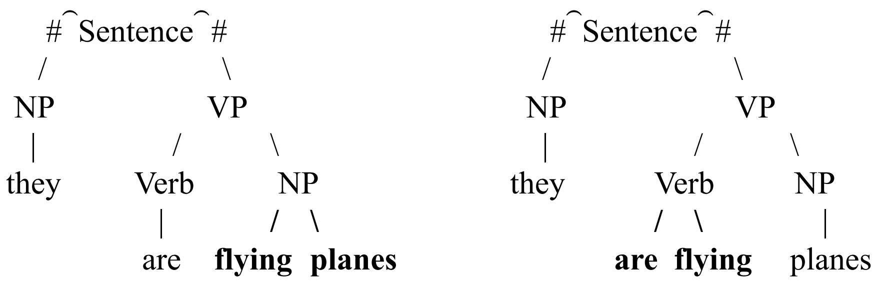
Noun Phrase → they

Noun Phrase → planes

Noun Phrase → flying⁀planes

This set of rules in (4.3), according to Chomsky (ibid), provides two “non-equivalent derivations of the sentence” in question, as represented in two different tree diagrams. Below, in (4.4), I replicate the diagrams from the original paper but highlight the parts that make them distinctive from each other.

(4.4)



These two tree diagrams in (4.4) capture two understandings of the sentence: either as ‘they - are - flying planes’ or ‘they - are flying - planes.’

Generative linguistics is often considered as neglecting semantics (e.g., Putnam 1961). While this assessment may be correct about Chomsky’s overall research program (which is out of the scope of this dissertation to judge), semantics does seem to play a role in the conception of a phrase structure grammar. For instance, ambiguity and its resolution can be seen as a testbed of the ‘empirical adequacy’ of a grammar. When a proposed grammar provides multiple trees for some sentence, it is said to be “a case of constructional homonymity” (1956, 118, emphasis original). Constructional homonymity is a formal property, but Chomsky suggests that it can be used “as an explanation for the semantic ambiguity of the sentence in question” (ibid). Conversely, he says, one may reject “the adequacy of a grammar ... by determining whether or not the cases of constructional homonymity are actually cases of semantic ambiguity” (ibid).

4.2.2 Formalizing ‘Immediate Constituent Analysis’

Sheila Greibach (1981, 15) argues that PSG is one attempt to formalize the method “used by linguists in describing the morphology and syntax of natural languages” called immediate constituent analysis. Immediate constituent analysis resembles the analysis in propositional logic where compound sentences are analyzed into atomic sentences with logical operators connecting them. In linguistics, the analysis goes all the way to the morphemes, the smallest, meaningful elements of a language.[31] A sentence is divided into two or more immediate constituents ('ICs'), "each IC into ICs, and so on down" (Greibach 1981, 16). For example, the sentence 'the king of England opened Parliament' can be analyzed as follows (Wells 1947, reviewed in Greibach 1981):

(4.5)

the || king ||| of |||| England | open||ed || Parliament

The string in (4.5) indicates that two ICs 'the king of England' (the subject) and 'opened Parliament' (the predicate) are analyzed in the first step. The first IC is then analyzed into 'the' and 'king of England,' whereas the second IC 'open', '-ed', 'Parliament.' Then 'king of England' is further analyzed into 'king,' 'of,' 'England.' This process stops once it reaches individual morphemes.

According to Greibach (1981), Zellig Harris, Chomsky's dissertation advisor, made one of the first attempts to formalize the process of IC analysis. In *From Morpheme to Utterance*, Harris (1946) introduced an equation of some sort (e.g., $BC = A$) to describe the formation of utterance from the level of the morphemes up. He first defines classes of morphemes or sequences of morphemes that can appear in the same context in the language. Then he forms equations to describe the replacement relationship. $BC = A$ states that "a morpheme of class B followed by a morpheme of class C can be substituted for a morpheme of class A" (Greibach 1981, 16). As an utterance would be divided into morphemes that are placed into different morpheme

---

[31] For example, the word *incoming* can be analyzed into three morphemes: *in*, *come*, and *-ing*.

classes (e.g., words, sentences), the equations were meant to be used to make repeated substitutions until the whole utterance is grouped into a sentence type, i.e., a morpheme class that corresponds to a whole sentence (Harris 1946, reviewed in Greibach 1981).

Coming back to Chomsky's PSG, his formal analysis of a sentence was indeed presented in the context of the IC analysis. In his words (1956, 136),

> Customarily, syntactic description is given in terms of what is called "immediate constituent analysis." In description of this sort the words of a sentence are grouped into phrases, these are grouped into smaller constituent phrases and so on, until the ultimate constituents (generally morphemes[32]) are reached. These phrases are then classified as noun phrases (NP), verb phrases (VP), etc. For example, the sentence ['the man took the book'] might be analyzed as in the accompanying diagram [see Fig. 4.4].

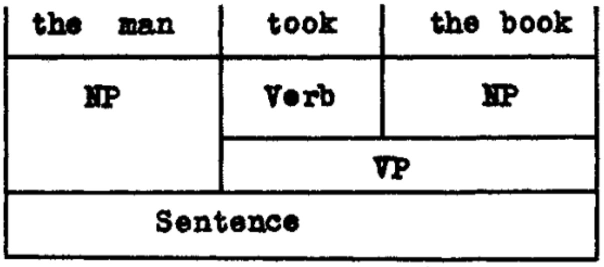


Fig. 4.4 A diagram of Phrase Structure Analysis. As shown in Chomsky (1956, 117), this diagram demonstrates the different categories that may be used to group different constituents within a sentence based on phrase structure grammar.

Formalizing IC analysis was crucial to the development of two 'machine translation' programs. The first machine translation program, which I will discuss in

---

[32] The author notes that "By 'morphemes' we refer to the smallest grammatically functioning elements of the language, e.g., "boy", "run", "ing" in "running", "s" in "books", etc. (Chomsky 1956, 124)

Section 4.4, refers to a study of the automatic translation between computer programs and the machine language. In contrast, the second machine translation program refers to a study of automatic translation between natural languages, which is the predecessor of today's computational linguistics. In 1950s, the objective of machine translation was to develop an automated mechanism to translate texts in Russian language into English without human intervention (O'Regan 2013). The task turned out to be considerably more complex than expected. Later, the focus shifted to develop algorithms and software for processing natural languages, eventually leading to a sub-discipline now call 'computational linguistics' (O'Regan 2013, Poibeau 2017).

4.2.3 The Correspondence Hypothesis: the Models 'for' Parsing Linguistic Information vs. the Models 'of' Utterance Production and Comprehension

Computational linguistics is often associated with the study of artificial intelligence (O'Regan 2013) and psycholinguistics (Schubert 2019). This association is consistent with the metaphysical commitment or methodological assumption some researchers held: Namely, that the mental processes of language production and comprehension can be understood in terms of information processing, which in turn can be understood through studying syntax (e.g., Lidz 2018, also see Hornstein et al. 2018). However, at the time when formulating IC analysis was an active research program, not all linguists agree on what IC analysis *described*.

On the one hand, the equations emerging from IC analysis were interpreted by some linguists as merely a means for parsing. In other words, those formal systems were models *for* parsing linguistic information. For instance, Wells (1947, 100, quoted in Greibach 1981) said that "The task of IC-analysis is the task not of describing what utterances occur but of describing, after these utterances have been given, what their constituents are."

On the other hand, some other linguist(s) (e.g., Harris 1946) understood those formal systems to be both models *for* parsing linguistic information and models *of* the information processing related to utterance production and comprehension.[33] This latter view is sometimes referred to as the correspondence hypothesis, i.e., there is a correspondence between the structure (e.g., NP, VP) set up by the linguist and the structure used by the speaker or hearer in speech production or comprehension. In other words, linguistic structures are psychologically real (Harris 2010).

Unlike his thesis advisor, Chomsky did not seem to (fully) endorse the correspondence hypothesis. His stance regarding the hypothesis can be seen in a footnote in which he defended his terminology. The term 'generate' in generative linguistics, which Chomsky borrowed from Post (1944), has been a focus for criticism of Chomsky's study of natural languages. As he wrote (1959b, 137-8), such a concept of "sentence-generating grammars" has "erroneously been interpreted as" only taking into consideration "language from the point of view of the speaker rather than the hearer." Actually, it does neither; Chomsky continues (ibid., 138):

> [S]uch grammars take a completely neutral point of view. ... We can consider a grammar of *L* to be a function mapping the integers onto *L*, [the] order of enumeration being immaterial (and easily specifiable, in many ways) to this purely syntactic study, though the question of the particular "inputs" required to

---

[33] To be fair, Harris did warn that "there are further limitations of selection among the morphemes, so that not all the sequences provided by the formulae occur" in actual language-use (Harris 1946, 178, quoted in Greibach 1981).

> produce a particular sentence may be of great interest for other investigations which can build on syntactic work of this more restricted kind.

This quotation suggests that generative grammars were not constructed to capture the processes when utterance production or comprehension is in action (i.e., 'linguistic performance'). That is, similar to the stance Wells (1947) had taken, neither the derivation of a string nor the rules of production was meant to be interpreted as models *of* actual information processing of speech production or comprehension, be they psychological or neurological.

That said, as shown in the second half of the quotation above, where the question of 'inputs' comes in, Chomsky does not rule out the utility of his 'purely syntactic study' in the empirical approaches to explain linguistic behavior. As a matter of fact, psychologists including George Miller have attempted to investigate the 'psychological reality' of linguistic transformations (reported in Harris 2010, Miller 1967, Baars 1986, Gardner 1985). The inquiry in psycholinguistics in the 1960s and the 1970s, which was based largely on the conjecture "that the more transformations differentiated two sentences, the longer it would take for people to relate them," did not pan out (Harris, 246, also see Gardner 1985).[34] Nonetheless, one product of this research program is an experimental protocol for probing the subject's ability to learn artificial grammars. It is this experimental protocol, now referred to as 'artificial grammar learning' (AGL), that cognitive biologists apply to test the supra-regular hypothesis, which I discuss in Chapter 5. In the section that follows, I discuss Chomsky's argument for the 'supra-regular

---

[34] According to the linguist Stanley Dubinsky at University of South Carolina (in personal conversation), a similar assumption remains in today's empirical linguistics.

thesis,’ which is the basis of the supra-regular *hypothesis*.

4.2.4 Modeling The Syntax Of Natural Languages: Chomsky’s Argument for the Supra-regular Thesis

Chomsky (1956) argues that the grammars of the most restricted class in the Chomsky hierarchy are too limited to be considered an adequate model for the syntax of natural languages. His argument can be viewed as intended to guide researchers to a plausible method for modeling the syntax of natural languages. I reconstruct his argument (as in Chomsky 1956) and highlight the central role of the hierarchy in **P1** of the argument shown in (4.6).

(4.6)

P1: If there exists a robust syntactic feature in at least one natural language that is indescribable by a given class of formal grammars in the Chomsky hierarchy, then it requires formal grammars beyond that particular class in the hierarchy to model the syntax of natural languages. (‘*The Classificatory Conditional*’)

P2: There exists a robust syntactic feature in English indescribable by finite-state grammars (i.e., the most restricted class in the hierarchy) called the ‘long distance dependency’.[35] (‘*The case of the long distance dependency in English*’)

C: It requires a formal grammar beyond the most restricted class in the Chomsky hierarchy to model the syntax of all natural languages. (‘*The supra-regular thesis*’)

The argument in (4.6) has the form of Modus Ponens in which the Chomsky hierarchy is featured in the consequent of **P1** and, consequently, in **C**. In particular, **P1**—‘the

---

[35] According to Larson (2017), Harris (1945) described a similar feature but termed it ‘discontinuous morphemes’.

classificatory conditional'—can be viewed as a methodological statement, declaring the criterion of when to 'move up' the hierarchy in search for the right class of formal grammars to model natural languages.

Insofar as the argument in (4.6) is sound, the crux of it lies in **P2**, the syntactic feature of the long distance dependency in English. To illustrate, consider the three sentence patterns in (4.7).

(4.7)

(i) If S_1, then S_2.

(ii) Either S_3, or S_4.

(iii) The man who said that S_5, is arriving today.

According to Chomsky (1956, 115) "these sentences have dependencies between 'if'-'then,' 'either'-'or,' 'man'-'is.' " For instance, replacing 'if' in (4.7i) by 'either' requires a corresponding replacement of 'then' by 'or,' or else, the resulting string would not be a sentence of English. Similarly, replacing 'man' in (4.7iii) by 'men' also requires a corresponding replacement of 'is' by 'are.'[36] In other words, the feature of long distance dependency is concerned with the dependent constituents—within a single sentence with regard to a language—that are not adjacent to each other. Note that such dependent constituents, if they occur in a sentence, occur in pairs. Thus, let m represent the degree of long distance dependency within any given sentence, S, in English. Then, m of S can be determined by counting the number of paired dependent constituents within S. For example, assume $m = 0$ for S_1 through S_5 in (4.7). Then for the sentence patterns

---

[36] The pair 'men'-'are' was not mentioned in Chomsky (1956).

from (4.7i) through (4.7iii), $m = 1$.

Moreover, Chomsky continues (1956, 115), in English, “there are infinite sets of sentences ... with more than any fixed number of” such dependent pairs of constituents. Consider the sentence patterns in (4.8) for instance.

(4.8)

(i) If *if* S_1*, then* S_2, then S_2.

(ii) Either *either* S_3*, or* S_4, or S_4.

(iii) The man who said that *the man who said that S5, is arriving today* is arriving today.

In (4.8), each of the italicized sequences indicates the replacement of S_1 in (4.7i) by (4.7i) itself, S_3 by (4.7ii) itself, and S_5 by (4.7iii) itself. After these replacements, the degree of long distance dependency increases from $m = 1$ in (4.7i-iii) to $m = 2$ in (4.8i-iii). If one were to repeat this self-replacement indefinitely, each replacement will result in a string that remains grammatical in English with an increased *m*.

Furthermore, Chomsky (1956) argues that a finite-state grammar cannot describe a language wherein the sentences do not have a fixed *m*. “Specifically,” he says (ibid., 114, emphasis original):

> we define a <u>finite</u>-<u>state</u> <u>grammar</u> G as a system with a finite number of states S_0, .., S_q, a set $A = \{ a_{ijk} \mid 0 \leq i, j \leq q; 1 \leq k \leq N_{ij}$ for each $i, j \}$ of transition symbols, and a set $C = \{ (S_i, S_j) \}$ of certain pairs of states of G that are said to be <u>connected</u>. As the system moves from state S_i to S_j, it produces a symbol $a_{ijk} \in A$.

See Fig. 4.5 for an illustration in Chomsky (1956).

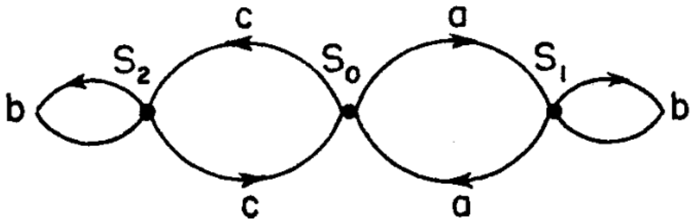


Fig. 4.5 A Diagram of a Three-State Finite-State Grammar

Shown in Fig. 4.5 is a "three-state process with (S_0, S_1), (S_1, S_1), (S_1, S_0), (S_0, S_2), (S_2, S_2), (S_2, S_0) as its only connected states, and with a, b, a, c, b, c as the respective transition symbols" (Chomsky 1956, 116). Thus, "this process can produce the sentences" such as "$a^\frown a$, $a^\frown b^\frown b^\frown a$, $a^\frown b^\frown b^\frown b^\frown a$, …, $c^\frown c$, $c^\frown b^\frown c$, $c^\frown b^\frown b^\frown a$, etc" (ibid.). According to Chomsky (ibid., 116), the language thus generated "has sentences with dependencies of any finite length."[37]

Also according to him, (ibid., 115), if a sentence in a language L has *m* pairs of nonadjacent, dependent constituents, then "at least 2^m states are necessary in the finite-state grammar, G, that generates the language L." Note that, a finite number of states S_0, .., S_q is part of the definition of a given finite-state grammar G. In other words, any given G has a definite number for q, e.g., when q = 3 with regard to G, G has 3 internal states, S_0, .., S_2, as illustrated in Fig. 4.5. Consequently, the degree of long distance dependency, *m*, in the sentences generated by a finite-state grammar defined by Chomsky is 'fixed' — for *m* in this case is determined by q, the number of possible finite states. It follows that—the conclusion—no finite-state grammars can describe the sentences in English where the degree of long distance dependency, *m*, is indefinite.

Finally, that conclusion entails **P2**, i.e., there exists a robust syntactic feature in

---

[37] There does not seem to further explanation for this statement in Chomsky (1956).

English indescribable by finite-state grammars. Because **P2** is the antecedent of **P1**, by Modus Ponens, it follows that **C**: to model the syntax of English, it requires a less restrictive class of grammars than the finite-state grammars. Recall that finite-state grammars are also called 'regular grammars.' Any classes of grammars beyond the class of regular grammars are 'supra regular.' To distinguish this conclusion in **C** from a hypothesis about the infrastructure of human mind, I refer to **C** as the 'supra-regular thesis.'[38]

4.2.5 The Supra-regular Hypothesis

Chomsky's supra-regular thesis was about the syntax of natural languages, but it has been interpreted as a hypothesis about human cognition (e.g., Fitch 2014). One dominant theme in the study of information in the 1950s was that the brain is a hardware that processes information, whereas the mind is the program that runs on the brain (e.g., Miller 1967, Gardner 1985). Moreover, according to the Chomsky hierarchy, finite-state grammars are equivalent to finite-state automata.[39] Thus, Chomsky's supra-regular thesis is taken to license the inference as follows (e.g., Fitch and Friederici 2012, Fitch 2014):

The Supra-regular Hypothesis — In order to process natural languages that are

---

[38] Pullum (2011, 277) has argued that "it is not clear that Chomsky ever gave a sound mathematical argument for" the claim that "English is beyond the power of finite state description". Indeed, Chomsky's (1956) original presentation of the argument is neither a mathematical proof nor a neatly organized argument. My reconstruction of that argument in this section is meant to be a charitable reading of the text instead of a challenge to Pullum's assessment.

[39] For any language L_1 that can be generated by a finite-state grammar G, one can construct a finite-state automaton A such that L_1 can be recognized by A, and vice versa, i.e., for any language L_2 that can be recognized by A, one can construct a finite-state grammar to generate L_2; it can be proved that $L_1 = L_2$, and thus, G and A are equivalent.

> supra-regular, the human mind must be the sort of program that is less restricted than the regular grammar. For the same reason, human brains must be less restricted than the finite-state automata to run the program that is supra-regular.

In other words, the supra-regular hypothesis is a product of both Chomsky's study of natural languages and the overall theoretical commitment to view the mind and the brain as a program that processes information and its physical implementation, and to approach them accordingly.

Of course, the brain is but one of such physical implementations. In the next section, I turn to discuss FLT in computer science, a discipline that studies information processing by building both the programs and their physical implementations, i.e., the software and the hardware, respectively.

4.3 Theory of Formal Languages in Computer Science

In computer science, what a formal grammar does is also to describe a formal language; however, the purpose for doing so is different between linguistics and computer science. In generative linguistics, the initial goal was to *reveal* 'the one grammar' that describes presumably all natural languages (Chomsky 1956). That is, a grammar with the rewriting rules that can be used to generate the set of all and only sentences in each and every natural language. Moreover, regardless of whether such a grammar can be found, the derivations of sentences based these rewrite rules have provided a means for scholars to *theorize* the psychology of information processing (e.g., Lewis and Phillips 2015). In computer science, however, the purpose of describing a language (i.e., a programming language) is not so much about *revealing* 'the real grammar' nor is it about *theorizing* how information might be processed in the computer. Instead, as I argue in 4.3.1 and 4.3.2, the purpose is to *engineer* the flow of information processes in the computer and to *standardize* such engineering practices that we call programming (Ledley 1962). In 4.3.3, I illustrate my point by the development of the programming language ALGOL as an example. In 4.3.4, I discuss the use the Chomsky hierarchy as a classification scheme in computer science, using the discovery that ALGOL is context-free as an example. Finally, in 4.3.5, I argue that although both linguistics and computer science use the Chomsky hierarchy to classify languages, there is a crucial difference between the two classificatory uses.

4.3.1 Engineering the Flow of Information Processing

In computer science, the flow of information within a hardware computer is a result of engineering. For instance, Ledley (1962, 6, emphasis original) introduces computers as "numerical-transformation" machines, where "numbers are the *inputs* to it, and the computer transforms these numbers into new numbers, which appear as the *outputs*." See Fig. 4.6.

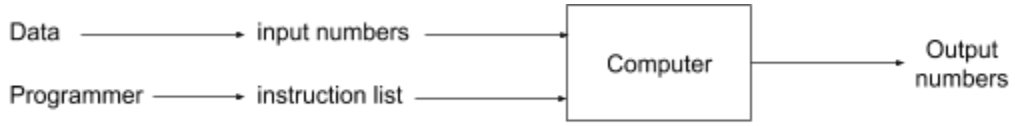


Fig. 4.6. Data Flow to Computer.

This diagram in Fig. 4.6 is a duplicate of Ledley's (1962, 7) diagram captioned "Data flow to computer." Note that the input to a computer (to the left of the box) consists of two different types, the input numbers and a list of instructions, both also in the form of numbers.

To illustrate, the input numbers may be the initial conditions of a differential equation or the raw experimental data, and the output numbers will be, respective to the input, a table of the functional solutions or the parameters determined by the experiment, and so forth (Ledley 1962). Moreover, in order for the computer to deliver expected output with respect to an input, it needs a different type of input numbers: *an instruction list*. Such a list is prepared and provided by the software engineers to automate the rewriting of numbers within the computer. In Ledley's words (1962, 7, emphasis original):

> [T]he computer must be directed to perform the required computations, and hence, in addition to the input numerical data, there must be also input instructions, which are coded in the form of numbers. Such a coded list of instructions is called a *program*, or code; the preparation of this program or code is the responsibility of the computer *programmer*.

In other words, by providing the instruction list as input numbers, what a computer programmer does is to engineer the transformation process between the *data* as input and the expected numbers as output, i.e., the flow of information processing within the computer. This engineering use of FLT distinguishes the applications of FLT in computer science from Chomsky's application of FLT in linguistics, a point I will revisit in Section 4.4

4.3.2 Standardizing the Practice of Programming

Early instruction code was written in machine language, with sequences of 0's and 1's. The computer performs certain arithmetic operations by following the machine code to flip the switches inside the computer on or off. The 1950s saw the development of higher-level programming languages, which allowed programmers to specify instructions using mnemonic codes (abbreviations of operations, e.g., ‘inc’ means ‘increase by one’) and with constructs such as loops, arrays, and procedures. Before a hardware computer could perform the expected operations, these languages need to be converted (i.e., compiled) into executable machine language through an assembler.

Before the 1960s, programming languages were not ‘portable.’ Not all programming languages were compatible with all of the available hardware computers. For example, the code written in the programming language FORTRAN, which was developed by IBM in the late 1950s and the only mainstream programming language at that time, could only be compiled and run on IBM’s machines (Savage 1998). Consequently, changing between different hardware or upgrading the hardware meant rewriting programs. To improve portability of the programs, it became important to establish “a universally acceptable international automatic language, analogous to the universal written language of music” (Ledley 1962, 203). In 1960, after a series of conferences and meetings among representatives of multiple European countries and the United States, the first of such universal automatic languages, it came the Algorithmic Language ‘ALGOL’ (Backus et al. 1960). The meta-language developed to specify the syntax of ALGOL became known as Backus-Naur Form (BNF) and was used to standardize the syntax of other programming languages (Hyman 2010).

4.3.3 Backus-Naur Form and ALGOL

In addition to being the first language established for standardizing the programming practice, ALGOL is also the first programming language to be constructed in the manner that resembles PSG as Chomsky constructed in linguistics (Ledley 1962). That is, PSG to a formal language in linguistics is as BNF to ALGOL in computer science. Indeed, it is recognized that while “BNF was developed independently by Backus, it is really one of Chomsky’s grammars with a different notation” (Sammet 1972, 607; also see Moll, Arbib, and Kfoury 1988, 2). To illustrate, in what follows, I use BNF to describe a few examples of the syntax of ALGOL. This illustration is to show the components of ALGOL using BNF, the relation between these components, and how they resemble the way Chomsky describes formal languages using PSG. In the illustration of syntax, however, I will also point out ‘recursion’ rules as I go; the topic of recursion will be of focus in Section 6.4.

ALGOL is composed of ‘characters,’ ‘expressions,’ and ‘statements,’ which are roughly analogous to ‘symbols,’ ‘ strings,’ and ‘sentences,’ respectively. However, because the purpose of ALGOL is to describe processes of numerical transformations, ALGOL has a more elaborated set of characters than the alphabet of PSG. For example, consider the characters in (4.9).

(4.9)

<letter> ::= **a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z**
|A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z

<digit> ::= **0|1|2|3|4|5|6|7|8|9**

<logical value> ::= **T** | **F**

<arithmetic operator> ::= + | − | × | ⁄ | ↑ |

<relational operator> ::= < | ≤ | = | ≥ | > | ≠

<logical operator> ::= **⊃** | **∨** | **∧** | **¬** | **=**

<delimiter> ::= **(** | **)** | **[** | **]** | **:** | **.** | **;** | **,** | **:=**

In (4.9), '<______>' indicates categorical rules that define 'the structure ______ of the language'. The characters of ALGOL consist of seven different type of symbols, as indicated between the angle brackets. '::=' represent 'is defined as'. The symbols on the right hand side of '::=' are not in brackets because they are in the language of ALGOL, whereas those on the left part are in BNF. Moreover, most of the operator symbols assume their common interpretations. The arithmetic operator '↑' means 'raise to the power,' e.g., A^2 would be written 'A↑2'. The delimiter '**:=**' roughly means 'replaces,' i.e., the left part is to be replaced by the name (or address) of the value of the right hand side, e.g., '**A := 3**' is a statement that directs the computer to assign the value of **3** to **A**. (Ledley 1962, 218).

To see some syntactic rules of ALGOL and how they work, consider the definitions in (4.10) and (4.11) for producing the expression types of 'numbers' and 'simple variable,' respectively.

(4.10)

i. <unsigned integer> ::= <digit> | <unsigned integer> <digit>

where the juxtaposition of two angle brackets '<...> <...>' means 'followed by'.

ii. <decimal fraction> ::= **.**<unsigned integer>

iii. <decimal number> ::= <unsigned integer> | <decimal fraction>
| <unsigned integer> <decimal fraction>

iv. <number> ::= <decimal number> | **+**<decimal number>
| **−**<decimal number>

(4.10i) states that 'an unsigned integer' can be either a digit or an unsigned integer followed by a digit. Note that this syntactic definition is called 'recursive' because syntactically, '<unsigned integer>' appears on both sides of the definition sign. In general, a definition is recursive when at least one of the terms it defines is used in the definition. (4.10ii) states that a delimiter '**.**' followed by an unsigned integer is a decimal fraction. In (4.10iii), it states that a decimal number can be expressed in one of the three forms: an unsigned integer (as defined in *i*), a decimal fraction (as defined in *ii*), or an unsigned integer followed by a decimal fraction (as defined in *iii*). Finally, in (4.10iv), it states that a number can be expressed in also one of the three forms, namely, a decimal number, a plus sign followed a decimal number, or a minus sign followed by a decimal number.

(4.11)

<simple variable> ::= <letter> | <simple variable> <letter>
| <simple variable> <digit>

(4.11) states that a simple variable is a letter, a simple variable followed by a letter, or a simple variable followed by a digit. Note that this definition is also recursive, and based on recursive reasoning, the leftmost character of a simple variable is always a letter.

4.3.4 ALGOL is Classified as a Context-Free Language

Both theoretical and practical, the significance of the Chomsky hierarchy to computer science started to show after Ginsburg and Rice (1962) proved that ALGOL was equivalent to a context-free, or Type 2, language (Hyman 2010).[40] Theoretically,

---

[40] According to the Chomsky hierarchy, a grammar that includes at least one recursive definition is supra regular; this is a direct consequence of how the restrictions are placed at the Type 3 class of grammars.

Ginsburg and Rice's discovery demonstrates that "BNF could express the grammar of all context-free languages and was equivalent to one form of Chomsky's phrase structure rules" (Hyman 2010, 270). As Hyman (2010) recounts, subsequent developments include the algebraic formulations for these formal languages (Shützenberger 1961; Chomsky and Shützenberger 1963) and eventually the proof that each of the languages on the Chomsky hierarchy corresponds to a particular type of automata (Kuroda 1964). Within fifteen years of Chomsky's 1956 essay, "no serious study of computer science would be complete without a knowledge of the techniques and results from language and automata theory (Hopcroft and Ullman 1969, v). Finally, the result of these developments was the new field of study, aptly named 'formal language theory,' that emerged from both the study of languages and the study of automata (Ginsburg 1980).

Practically, Ginsburg and Rice's (1962) discovery led to 'syntax-directed' compiler design (Hopcroft and Ullman 1969). To illustrate, in software engineering, programmers write code using particular formal languages, such as Python, or C++. Before the hardware of a computer executes the code, there are two preliminary tasks: it needs to check whether that particular code belongs to the programming language in question, and then, if the answer is 'yes,' it needs to translate the code from that programming language to the machine language of the CPU. A compiler is itself a program that carries out these two tasks. To ensure performance, the writer of a compiler will need to make sure that the compiler he or she writes will always *reject* code that violates the rules of the programming language (i.e., the *syntactically incorrect* programs), as well as always accepting the syntactically correct ones — this part of a compiler is also called a 'decision program' - i.e., an automaton.

4.3.5 The Classificatory Use of the Chomsky Hierarchy

The Chomsky hierarchy helps compiler writers in designing compiler programs; it helps programmers to first classify the code that their decision program will be dealing with and then resort to the right template accordingly to engineer the program. Specifically, the fact that a programming language, *P*, belongs to a particular class of languages on the Chomsky hierarchy, *L*, entails that all the sentences in *P* (i.e., any code written in *P*) are describable by the class of grammars *G* that corresponds to *L*. That is, for instance, all the code written in ALGOL is describable by context-free grammars. Note that each *G* has an equivalent class of automata, *A*, e.g., the corresponding automata for context-free grammars are pushdown automata. Thus, (1) any code written in *P* is describable by *G* and (2) *A* is equivalent to *G*; together, (1) and (2) entail that *P* is recognizable by *A*, in principle. That is to say, knowing that ALGOL is context-free allows a designer to simplify their approach to engineer the compiler. The Chomsky hierarchy gives the designer a ‘template’ to work with, i.e., by incorporating a pushdown stack in the decision program to process ALGOL code.

Note that in computer science, compiler design is not the only area of study benefited from the Chomsky hierarchy. In general, any programmer whose task includes parsing input could take advantage from knowing how to properly *identify* the language and classify it accordingly. Consider the HTML code as input. Suppose that the set of all HTML code is supra-regular. Suppose also that ‘regular expressions’ are, as the name suggests, regular. Then it follows that the attempt to parse a HTML code using regular expressions in the decision program would be futile. However, there are exceptions.

Suppose that the set composed of all the telephone numbers in the U.S. is regular. Then, if all one needs from the HTML code is all and only the U.S. phone numbers, nothing else, then regular expressions would suffice. In other words, while parsing all possible HTML code is a 'supra-regular' problem, parsing only a particular kind of information in the HTML code may not be a supra-regular problem. Overall, having the Chomsky hierarchy as a classification scheme is useful to computer programming engineers — if the programming task at hand is not supra-regular, it does not require a supra-regular solution.

4.4 A Tool-migration Analysis

My discussion of FLT thus far allows for me to answer the two questions that I purported at the end of Chapter 3: What does it look like when the target profile of a tool in migration remains the same? Can a tool's usage profile change without the change of its target profile? In other words, what does tool-adaptation look like? I argue that the cross-disciplinary uses of the Chomsky hierarchy between linguistics and computer science discussed thus far presents such an example.

First, FLT including the Chomsky hierarchy can be analyzed as a research tool. As I characterized it in Chapter 2, a research tool is a piece of mathematical construct that guides its users to (1) formulate a *problem*, (2) obtain an *analytic result* based on the formulated problem, and (3) interpret the result as a *solution* to the initial problem. I recast my discussion in Section 4.2.4 in terms of 'problem,' 'analytic result,' and 'solution' as follows.

In Chomsky's initial application of FLT:

- the ‘problem’ is concerned with which type of models should be used to describe the syntax of natural languages,
- the ‘analytic result’ is the argument (and an alleged proof) that the feature of long-distance dependency in English is undescribable by finite-state grammars, and
- the ‘solution’ is stated in the supra-regular thesis, which is an interpretation of the analytic result, i.e., modeling the syntax of natural languages requires a supra-regular grammar.

In a sense, should there be an explicit rule in computer science — ‘if a problem is supra regular, then go for a supra regular solution’ — Chomsky should be given credit for its origin.

Second, the cross-disciplinary use of FLT between linguistics and computer science is a case of tool migration. In Chapter 2, I suggest that when a research tool is used to study a new subject matter (i.e., one that is different from its previous applications), then we have a case of tool migration. In Section 4.2, I have shown that FLT has been used to theorize information processing with regard to language. By doing so, it was hoped to explain linguistic phenomena, such as ambiguity, and to reveal the syntax of natural languages. What exactly is this ‘syntax of natural languages’? Chomsky (1965, 3) later elaborates it as “the speaker-hearer’s knowledge of his language.” Thus, one can say the subject matter of applying FLT in linguistics is the knowledge of a natural language, whatever it may mean. By contrast, as I have discussed using ALGOL as an example, the subject matter of applying FLT in computer science is to study the ways to develop, and to improve the engineering of, the flow of information processes within the computer. Thus, this contrast in the two subject matters licenses us to view the case of FLT between linguistics and computer science as tool migration.

Third, the next step is to determine which of the four types of tool migration better characterize the case at hand. To do so, there are two question that need answering: one is concerned with, in these two applications of FLT, the definitions of the entities in the formal system and the relations between these entities, namely, the tool's target profile; the other is concerned with the ways in which FLT is applied to operate in the two applications, i.e., the tool's usage profile.

In terms of the target profile, are there any changes in the assumptions or definitions regarding the 'alphabet,' 'phrases,' 'sentences,' 'rules,' 'derivations,' etc., when considering FLT in Chomsky's study of natural languages and the development of ALGOL? The answer to this question, I argue, is negative. In both linguistics and computer science, these theoretical terms in FLT are about the units of information and the relation between these units. In linguistics, symbols in an alphabet are the smallest units that form phrases, and phrases, in turn, are units that form sentences. In the design of ALGOL, characters are the smallest units that which form expressions, and expressions, in turn, are units that which form statements. In both disciplines, the grammatical rules determine the derivation of these units. These crucial similarities suggest that the target profile of FLT has survived through the migration.

In terms of the usage profile, are there any changes in the way in which FLT is used between the two applications I looked at thus far? The answer to this question, I argue, is positive. In linguistics, FLT has been used to (1) explain ambiguity, (2) formalize a means of analysis, and (3) model the syntax of natural languages through means of classification. In computer science, FLT maintains these usages while adding a fourth one. In computer science, (1') the derivations trees of FLT are used to detect ambiguity (Parkes 2002)[41], (2') the meta-language BNF is formalized to analyze programming languages, and finally (3') the Chomsky hierarchy is applied to 'model' the data as input (e.g., code to be parsed by a compiler or by any program) through means of classification. Recall Fig. 4.6 for an illustration of the relation between 'data as input,' 'instruction as input' and the 'output' of the computation. However, despite these similarities, the 'engineering use' of FLT in computer science is a salient, new way of applying FLT. In particular, in computer science, programmers prepare the list of instructions in order for the computer to produce intended output. The engineering aspect of using the Chomsky hierarchy, thus, stands out from the theorizing aspect of using the same tool in linguistics: the major distinction is that engineering is 'output oriented.' To illustrate, recall the task of parsing phone numbers in HTML code. In that task, the relevant data as input are the HTML code, which is supra-regular, whereas the intended output, i.e., the U.S. phone numbers in the code, are not. In a sense, the way programmers model the input in such a task is directed by the intended output, not necessarily the nature of the input. This output-oriented, engineering use of FLT differs from the theoretical use of FLT in linguistics that I discussed in Sections 4.2-4.3. This engineering use does not appear in the application of FLT when it migrates to experimental psychology, a story that I resume in the next chapter.

In sum, FLT provides a framework to think about engineering problems and

---

[41] In my discussion, I did not touch on the point in (1'), for the details in which computer scientists use the derivation trees to detect or explain ambiguity does not affect my argument.

solutions to tackle these problems. These problems are analyzed in terms of varying units of information and the relation between these units; the solutions are approached in terms of the processes that generate them and the processes that parse them. The main observation emerged from the discussion of this chapter is thus: In linguistics, FLT is used to theorize and account for linguistic phenomena, whereas in computer science, it is utilized to improve the implementation of information processes. This additional, engineering use of FLT in computer science augments and, consequently, changes the usage profile of the tool. Taken together, because the target profile survives the migration but not the usage profile, the case of FLT migrating from linguistics to computer science is, I conclude, an example of tool-adaptation.

Chapter 5

Formal Language Theory in the Sciences of Mind, Brain, and Information Processing

In this chapter, I switch to explore one positive impact brought by migrating the research tool of formal language theory ('FLT') to the sciences of mind and brain. Using nonhuman animals as model organisms for understanding human linguistic abilities was one signature of the behaviorist approach, once dominant in experimental psychology — ironically, Chomsky was widely regarded as the key figure for its decline (Orgood 1975, Gardner 1985). And yet, in this new wave of experimental psychology, songbirds and nonhuman primates re-appear in the quest for knowledge concerning human verbal behavior, all under the framework of FLT due to Chomsky. One may ask: How do scientists reconcile the differences between the behaviorist and the cognitivist approaches? In particular, how do they combine using animals as model organism and formal models of natural languages together to form a burgeoning research program? This chapter is an attempt to start to address this latter question from the perspective of tool migration. I assume that the answer to the latter question will inform an attempt to answer the former, even though a proper attempt is outside the scope of this dissertation. I will argue that the key that introduced this new, hybrid, strand of empirical psychology was a 'conceptual crossover' due to a similar target profile with an augmented usage profile. In this chapter, a case of tool-adaptation is shown to integrate animal model organisms and formal models of languages into one scientific framework, called *cognitive biology* (Fitch 2014).

To begin, in Section 5.1, I review an experiment in the 1960s in which George Miller introduced artificial grammars to experimental psychology. This experiment became a prototype of what is now known as the experiment of artificial grammar learning ('AGL'). In Section 5.2, I discuss the first reported experiment that incorporates the protocols of AGL with the Chomsky hierarchy. Fitch and Hauser (2004) design an experiment of AGL by using two grammars, a Type 3 grammar and a Type 2 grammar, respectively, in the Chomsky hierarchy. Unlike their predecessor, Fitch and Hauser interpret their experimental results based on the Chomsky hierarchy, in particular, the concept of supra regularity. By doing so, the duo is able to relate animal model organisms with formal models of languages in the study of the evolution of language. In Section 5.3, I argue that their achievement can be accounted for in terms a particular type of tool-adaptation: an augmented usage profile with a similar target profile.

5.1 Project Grammarama: A Prototype of The Artificial Grammar Learning Experiment Protocols

Known for his classic paper "The Magical Number Seven, Plus or Minus Two," Miller's other classic contribution to psychology is the prototype of the AGL experiment. The original idea was a guessing game involving an artificial language instead of natural languages like English. The idea gradually evolved into a research program, and the initial idea was implemented in various forms, but the goal has been the same—to understand rule learning. Miller (1967, 129 emphasis original) depicts his guessing game as follows:

> Suppose that someone is shown arbitrary strings of signals and allowed to ask which are grammatical and which are not. Suppose moreover that the rules of our game do *not* resemble English grammar, and that the signals are *not* familiar English words and phrases. In this ... form there is no limit to the variety of artificial grammars we could invent. Imagine, in short, a completely novel, completely abstract grammar and meaningless vocabulary. With nothing more than that—with no meanings, no sensible use of the strings, with nothing but formal criteria as a guide—can a person discover the grammatical rules underlying the language? And if so, what is the best way for him to get at it?

In the 1950s, Miller conceived of this game, which he named 'Grammarama.' Soon after computers were introduced to psychology laboratories, it became one of the first computerized experiments in the 1960s.

The initial Grammarama experiment is comprised of a learning phase and a test phase. The participant – Miller calls the participant 'the learner' – enters a small, soundproofed room equipped with a chair and a teletypewriter on the table. The experimenter communicates the instructions with the learner by printing them out on the teletypewriter. The learner then responds by using the keyboard. The typescript, thus, shows a record of all the exchanges during an experimental session.

A session starts with a description of the procedure: The learner is told to discover the rules for generating admissible strings of letters. He or she may type any strings composed from a specified alphabet. Upon completing a string, the experimenter prints out RIGHT to indicate a grammatical string, or WRONG to signal otherwise. The process will continue until the learner feels certain that he or she has learned the rules. By typing FINISH, the learner will enter the testing phase and be given in total 10 strings. For each string, the learner will type ‘C’ to indicate ‘correct’ (i.e., it’s a grammatical string) or ‘I’ to indicate otherwise. A score will then be printed to show the learner how he or she has performed, and the session ends.

Instead of the overall correctness of the test result, i.e., the end product of learning, the goal of Project Grammarama was to understand “how people learn the grammatical rules underlying artificial languages” (1967, 126). Unlike the ‘engineering use’ of FLT in computer science, the goal of the experiment shows a ‘theoretical use’ of FLT similar to that in linguistics. As Miller (1967, 164, emphasis mine) stresses:

> Interesting psychological processes occurred during the interaction; it is these processes that we hope to understand, *not their end product*. The question is, what cognitive functions do the protocols reveal?

At the time, 98 subjects were tested with regular grammar and context-free grammar, in each case with three sizes of the alphabet (two, three or six symbols). From the typescript of these 98 learners, a variety of learning strategies was identified as used by the subjects. It includes “anagrams, cycles, mirroring, counting, progressing, permuting, algorithmic, random, unique” (ibid., 165). For instance, “if x is admissible, then xx is also, and xxx, and, in general, x^n. This is what” Miller and colleague “called a cyclic strategy” (Miller 1967, 164). See (5.1) for illustration.

(5.1)

(2) DRDRDRDRD $D(RD)^4$

(3) RRDRRDRRD $(RRD)^3$

(4) RRRDDRRRDDDRRRDDRRRDD $(RRRDD)^4$

(5) DDDDDDDDDDR $(D)^{10}R$

In (5.1), listed in the left column in are the second through the fifth responses from one particular subject. Listed in the right column are notes made by the experimenter to indicate the pattern in each of the responses. Despite the goal was to understanding the process instead of the learning result, Miller noted that 54 out of their 98 subjects “typed the word FINISH and took the test before they were ready to pass it” (ibid., 167).

Details of these learning strategies are interesting in their own right but largely irrelevant to our present purpose. The primary reason to discuss this guessing game is to contrast it with Fitch and Hauser’s (2004) version of the AGL experiment. In Project Grammarama, the role of FLT and the Chomsky hierarchy is nearly inconsequential. Miller applied formal grammars to design artificial languages, but neither the grammars nor the Chomsky hierarchy was incorporated into his interpretation of the experimental results. Using the Chomsky hierarchy to design tasks that are either ‘regular’ or ‘supra regular’ would have to wait until a more recent wave of nonhuman animal AGL experiments.

5.2 Formal Language Theory in Cognitive Biology

5.2.1 First Experiment in Comparative Psychology where Formal Language Theory and Artificial Grammar Learning Join Force

The comparative psychologists Tecumseh Fitch and Marc Hauser (2004, 380) report experimenting with cotton-top tamarin monkeys (*Saguinus Oedipus*) using a regular grammars, $(AB)^n$, and a supra regular (context-free) grammar A^nB^n. In this now widely-cited experiment, Fitch and Hauser implement these two grammars by audio-recorded consonant-vowel syllables, spoken by a female and a male to create two sets of distinct acoustic elements, A and B. These two sets are differentiated in voice pitch (>1 octave difference), phonetic identity, average formant frequencies, etc.[42] For instance, Set A consists of syllables {**ba**, **di**, **yo**, **tu**, **la**, **mi**, **no**, **wu**} spoken by female, whereas Set B consists of syllables {pa, li, mo, nu, ka, bi, do, gu} spoken by male. Thus, strings such as '**no** li **ba** pa' and '**la** pa **wu** mo **no** li' conform to the $(AB)^n$ grammar as they are instances of $(AB)^2$ and $(AB)^3$, respectively. In contrast, strings such as '**yo la** pa do' and '**ba la tu** li pa ka' satisfy the A^nB^n grammar since the first string is an instance of A^2B^2, whereas the second string an instance of A^3B^3.

Fitch and Hauser divided twenty adult cotton-top tamarins into two groups, one per grammar, each group with a mixture of sexes and ages. Their experimental procedure includes a training phase, a re-familiarization phrase, and a test phase at the end. The training phase took place in the animals' home cages in one evening. During this phase, all subjects were simultaneously exposed to 20 minutes of repeated playback, which plays 60 different grammatical strings in random order.

The re-familiarization and the test phases took place in the next morning. When an individual subject wandered into a sound chamber, the experimenter played strings randomly chosen from those 60 training strings for two minutes while the animal was fed with treats. After this re-familiarization phase, the experimenter then closed the door of

---

[42] Examples of the sound files used in the experiment are provided by Fitch and Hauser and can be found online: http://science.sciencemag.org/content/suppl/2004/01/15/303.5656.377.DC1.

the sound chamber and started videotaping the subject. During the test phase, no food was delivered. When the subject was both looking down and away from the loudspeaker, the experimenter initiated playback of the test stimuli.

Based on the digitized video, multiple observers scored the latency and duration of looking (i.e., orientation towards the loudspeaker). The observers' scores agreed with one another >90% of the time. Between the groups, the testing stimuli are the same eight strings, four strings consistent with $(AB)^n$ and the other four consistent with A^nB^n. None of the eight strings were included in the training stimuli to either group of animals.

The reactions to the test stimuli between the two groups of monkeys were systematically different. In the group trained with the $(AB)^n$ grammar ('Group $(AB)^n$' henceforth), 9 out of the 10 tamarins looked more to the loudspeaker when violations were played. That is, they seemed to pay more attention to the strings generated by the A^nB^n grammar than those generated by their training grammar. Overall, in this group, the mean of looking to violations across animals is 72%, whereas the mean of looking to grammatically consistent novel stimuli is 34%. To Fitch and Hauser, this contrast between looking and not-looking behavior of the Group $(AB)^n$ suggests two things. First, the monkeys could distinguish between the two syllable-classes, Set A and Set B. Second, and more important, they are sensitive to their training grammar $(AB)^n$, in the sense that they might have learned the grammar.

This second point is a standard inference from the so-called 'familiarization-novelty' experiment protocols. It is common to infer whether a subject recognizes a pattern in the training stimuli based on the subject's reaction towards novel stimuli that lack such a pattern. The assumption is that subjects show more interest in novelty when they are sufficiently familiarized with the pattern in the training material. In Fitch and Hauser's experiment, all eight test strings are novel to the monkeys, but four of the strings share the same pattern as the training stimuli. The monkeys were trained with the $(AB)^n$ grammar. Thus, the expectation was that if they have learned the $(AB)^n$ grammar, they would look less to the loudspeaker when it played the strings consistent with the $(AB)^n$ grammar than they would when it played violations, i.e., strings of the A^nB^n grammar. In the authors' words, "the ability to learn the rule governing the construction of an acoustic sequence, without any explicit training, indicates that the tamarins are sensitive to regularities in an acoustic stream and can recognize novel strings as consistent with past inputs (2004, 379).[43]

In contrast, the group of subjects trained with the A^nB^n grammar ('Group A^nB^n' henceforth) showed no statistically significant difference in their looking behavior between the grammatical strings and the violations. Among the 10 monkeys, none looked at more than two of the four violations. Overall, the mean of looking to violations was 29% and the mean of looking to the consistent stimuli was 31%. Fitch and Hauser interpret tamarins' lack of 'special interest' in the $(AB)^n$ test strings as indicating that they failed to master the grammar A^nB^n, i.e., their training grammar.

5.2.2 Implications to the Evolution of Language

The purpose of Fitch and Hauser's AGL experiment is to probe the ability of tamarins to process acoustic sequences, but ultimately, what they are concerned with is the evolution of *human* linguistic capacity, especially the capacity to handle hierarchical structures. By

---

[43] However, this interpretation has been challenged by other psychologists, e.g., (Perruchet and Rey 2005).

‘hierarchical structures,’ the authors mean the structures such as ‘sentences,’ ‘phrases,’ ‘words,’ and ‘symbols.’ These structures are hierarchical because ‘sentences’ are composed of smaller units such as ‘phrases,’ ‘phrases’ are composed of even smaller units such as ‘words,’ and ‘words’ are composed of the smallest units such as ‘morphemes’ or ‘phonemes.’ In other words, the authors were intended to probe the ability of tamarins to recognize strings that could be generated by PSGs.

Their quest combines the interests from the disciplines of generative linguistics and evolutionary biology. As they put it, “linguistic syntax involves the rearrangement and permutation of ... abstract hierarchical structures” such as phrases or sentences in human languages that are above the word level (ibid., 377). On the one hand, the capacity in humans to perceive and produce these hierarchical syntactic structures “goes far beyond the simple concatenation procedures ... in animal communication” (ibid.). On the other hand:

> [T]he evolution of language faculty presumably involved the incorporation of some ancestral primate cognitive capacities. Thus, a critical question is whether hierarchical processing was one of these preexisting abilities. (ibid.)

In other words, for Fitch and Hauser (2004), despite lacking hierarchical structures in their communication systems (which may or may not be the case), some nonhuman primates may nonetheless possess the ability to process hierarchical structures. If indeed such an ability can be detected in other primates, which is the purpose of Fitch and Hauser’s AGL experiment, the finding will help shed light on how this ability have evolved within primate species. Even if the experiment result turns out to be negative, it may still inform neuroscientific investigation of linguistic abilities in humans.

That being said, it is crucial for Fitch and Hauser to experimentally demonstrate that humans indeed possess the said abilities, which they did. In the same paper (2004), Fitch and Hauser report their preliminary AGL experiment with adult human subjects using the same training stimuli previously mentioned in the tamarin experiment. According to the authors, all twenty adult human participants "showed rapid learning of either grammar (with under 3 min of exposure), and were easily able to discriminate grammatical from nongrammatical stimuli for both grammars" (2004, 379). From this stark contrast, especially the failure of the tamarins to recognize the strings generated by the A^nB^n (supra regular) grammar, they suggest that the acquisition of the hierarchical processing abilities "may have represented a critical juncture in the evolution of the human language faculty" (ibid., 380).

To narrow down among alternative explanations, the authors argue that the cause to tamarins' failure in learning the A^nB^n grammar must be *internal* to the subjects. According to them, all extraneous factors in the experiment were consistent between the two grammars. For example, the stimuli are of the same length and loudness. Subjects are able to perceive the A- and B-classes, as shown in their recognition of the $(AB)^n$ grammar. All of the duration of exposure, testing, or evaluation procedures were the same. Moreover, the authors argue, earlier work with this species using the same paradigm has demonstrated that these animals are capable of storing and recalling up to three separate stimuli and comparing them with subsequent strings. Therefore, they (ibid., 379) contend that the tamarins' inability to learn the supra-regular grammar, "does not result from some lower level limitation on memory, attention, or number discrimination." They acknowledge that it is possible that, with different experimental methods "(e.g., training and reinforcement), different grammars, and other species (e.g., apes)," other researchers might come to a different conclusion (379). Nonetheless, the authors conclude (380), "tamarins suffer from a specific and fundamental computational limitation on their ability to spontaneously recognize or remember hierarchically organized acoustic structures."

5.2.3 Relating Animal Model Organisms With the Study of Language Through the Chomsky Hierarchy

Fitch and Hauser (2004) appeal to FLT, especially Chomsky's supra-regular thesis, to interpret the experimental results. According to them, the stringset produced by the A^nB^n grammar requires a supra-regular grammar that "can embed strings within other strings, thus creating complex hierarchical structures ('phrase structures'), and long-distance dependencies" (378). Indeed, a discussion of this stringset can be found in Chomsky's (1956, Section 2) initial paper where he argues that finite-state grammars cannot describe long-distance dependencies in English. (See Section 4.2.4 for my reconstruction of Chomsky's argument for the supra-regular thesis). In that article, Chomsky gave three examples of formal languages that are also undescribable by finite-state grammars. One of the three, which I quote in (5.2), is likely to be the origin of this A^nB^n grammar in Fitch and Hauser's work.

(5.2)

> L_1 contains $a^\frown b$, $a^\frown a^\frown b^\frown b$, $a^\frown a^\frown a^\frown b^\frown b^\frown b$, ...,and in general, all sentences consisting of *n* occurrences of *a* followed by exactly *n* occurrences of *b*, and only these. (Chomsky 1956, 115)

Note that earlier in Section 4.1.2, I have already featured L_1 shown here in (5.2). In fact, it is a stock example of a supra-regular language. The grammar that generates it contains a set of two productions: $\{S \rightarrow aSb$ and $S \rightarrow ab\}$. Recall the three limiting conditions from the Chomsky hierarchy (Section 4.1.3). This grammar complies to all but the most restrictive condition, which I reiterate in (5.3):

(5.3)

> **Third limiting condition**: For every production $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ in P, (1) $|\boldsymbol{\alpha}| = 1$ and $\boldsymbol{\alpha} \in V_{N,}$ and (2) and $\boldsymbol{\beta}$ has the form $\boldsymbol{\gamma}$ or $\boldsymbol{\gamma\Psi}$, where $\boldsymbol{\gamma} \in V_T$, $\boldsymbol{\Psi} \in V_{N'}$ and $|\boldsymbol{\gamma}| = |\boldsymbol{\Psi}| = 1$, e.g., $A \rightarrow a$, $A \rightarrow aB$.

Specifically, this limiting condition in (5.3) does not allow for rewriting rules to have more than two symbols on the right-hand side. Since $S \rightarrow aSb$ in the set violates this condition, the whole grammar is classified as beyond Type 3, hence supra-regular.[44]

Fitch and Hauser (2004) also appeal to the tree diagram to represent the supra-regularity of the A^nB^n. To illustrate, consider the derivation of the string *aaabbb*: $\{S \Rightarrow aSb \Rightarrow aaSbba \Rightarrow aaSbbb \Rightarrow aaabbb\}$ (Section 4.1.2). This derivation is said to be represented in a tree diagram shown in Fig 5.1a. In Fig 5.1, I juxtaposition that tree diagram with Fig 5.1c, which Fitch and Hauser use to illustrate the hierarchical nature of the stringset generated by A^nB^n (2004, 378). In Fitch's later work (2014; Fitch and Federici 2012), he switches to talk of the 'tree-like hierarchical structures' as another way of referring to supra-regularity. That is, for instance, if a device can parse the stringset containing *aaabbb*, it must be able to process its derivation tree as shown in Fig 5.1a, and hence the device must be supra-regular. By contrast, according to Fitch and Hauser

---

[44] The grammar is a Type 2 context-free grammar.

(2004), the strings produced by the $(AB)^n$ grammar are sequentially organized. Such strings can be generated by a finite-state grammar ('FSG'), which makes it a regular language.

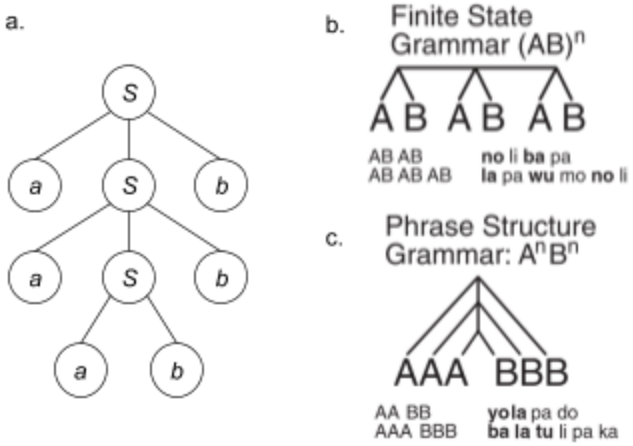


Fig. 5.1 Three Derivation Trees

In Fig. 5.1, on the left, it shows a derivation tree of the string *aaabbb*; on the right, it shows two diagrams in Fitch and Hauser (2004) to illustrate (bottom) the strings generated by a phrase structure context-free grammar A^nB^n as hierarchically organized and (top) the strings generated by a finite-state grammar $(AB)^n$ as sequentially organized.

With regard to the experimental design, the authors argue that the two stringsets generated by $(AB)^n$ and A^nB^n, respectively, make a useful pair of tools for investigating computational constraint. This is because the strings in both stringsets are formed from the same alphabet, i.e., the two syllable-classes, Set A and Set B. Thus, they argue, the only difference between these two stringsets is in the grammars that generate them. Consequently, if the tamarins can recognize the 'regular' stringset $(AB)^n$ but not the supra-regular stringset A^nB^n, then, as Fitch and Hauser reason, the ability of tamarins to recognize strings is limited at the regular level. Specifically, they say that (2004, 378):

> We chose the A^nB^n grammar because it is the simplest PSG that cannot, in principle, be approximated with an FSG but that can easily be brought into correspondence with a simple FSG in all nongrammatical respects, as required for our experiment.

Furthermore, they argue, the difference between FSG and PSG is more than of a matter of analysis. Their difference is 'real' both computationally and psychologically. In Fitch and Hauser's words (2004, 378):

> Like any PSG, the A^nB^n grammar requires additional computational machinery beyond a finite-state automaton. In computer science terminology, this addition would minimally be a push-down stack. In psychological terms, it requires some way to recognize a correspondence between either the groups formed by the As and Bs (e.g., counting) or between specific As and corresponding Bs (e.g., long-distance dependencies).

They highlight the experimental use of the grammar by concluding as follows (ibid.):

> This PSG thus provides the ideal grammar for the empirical issue addressed by this study by allowing us to focus on the generative power of the system without introducing extraneous performance variables (e.g., memory capacity or referentiality).

5.2.4 An Interdisciplinary Research Program was Born

Since its publication, Fitch and Hauser's (2004) experiment has brought a vital discussion to the scientific community concerning the biology and evolution of language. For instance, Donnell, Hauser, and Fitch. (2005, 286) argue that what FLT offers scientists (such as comparative, experimental psychologists) is a “more formal mathematical approaches to language.” Fitch (2010, 109) suggests that FLT is helpful for “formulating computational questions about how the human brain implements language, and for comparing human capabilities with those of other species.” Bowling (2014) also advocates that to understand the neural basis of human cognitive capacity, biologists need to first isolate those neural bases of cognitive capacities that are shared between human and nonhuman animals — the AGL experimental protocols powered by FLT provides just the right tool for this task. Fitch and Friederici (2012, 1940) argue that this renewed version of AGL shows “how FLT can be used, practically, by biologists, psychologists and neuroscientists, to design and execute experiments and analyse the resulting data.”

Emerging from these discussions is a proposal of an interdisciplinary research program, which features theory of computation as the hub to situate diverse domains of research such that their research results may inform one another. As Fitch argues (2014), for the science of cognition to mature, it needs to combine both the comparative and the computational approaches. In terms of the comparative approach, he urges cognitive scientists to expand their target interest from human cognition to include cognitive systems of nonhuman animals. In terms of the computational approach, he recommends the theory of computation for the role of bridging neuroscience, the cognitive sciences, and cognitive biology. As Fitch puts it (ibid., 330), neuroscience supplies “the firm physical foundations of brain function,” whereas cognitive biology provides “a comparative viewpoint.” Moreover, the link between disciplines “including cognitive and mathematical psychology, linguistics and musicology” should be built “upon the insights of computer science” (ibid., 330). Fitch concludes that the theory of computation and its “well-established branch called formal language theory” provides “an appropriate framework for ... considering high-level pattern perception of the sort typifying human ... language” (ibid., 347). By ‘high-level pattern perception,’ Fitch refers to supra-regularity, which, as he has argued, can be detected by the AGL experimental protocols that he and Hauser devised (2004).

5.2.5 The IFG-as-Stack Model

Over the course of ten years, the subject matter of the AGL research in cognitive biology grows from understanding the evolution of human language faculty (Fitch and Hauser 2004) to looking for the neural substrate of the pushdown stack. Based on these two general principles and the confirmation that human mind is supra regular, Fitch (2014) suggests that this neural substrate, once found, will explain how the human brain processes information that is ‘hierarchically organized,’ as illustrated by the derivation trees (Section 5.3.2). Fitch speculates that a likely location for the stack could be at the “inferior frontal gyrus (IFG, comprising Broca’s area and its neighbors), with sensory and association regions in the temporal and parietal lobes” (ibid., 355). IFG is the structure in the brain that “serves as a kind of ‘abstract scratchpad’ ” for offloading “partial results computed during serial processing of hierarchical structures” that are ensconced in the occipital and temporal cortices (ibid., 355). Calling this the “IFG-as-stack” model, Fitch (ibid., 355) establishes the link between neurolinguistics and computer science. In his words (ibid., 27):

> [R]everberations in the fronto-sensory feedback loop would play the role of the stack in the pushdown automaton implementing a context-free grammar. ... [T]he

> IFG would thus have an additional storage mechanism into which intermediate results (and in particular unfinished structural computations) could be placed for later retrieval.

Although Fitch admits that this IFG-as-stack model is highly speculative, he appeals to FLT as applied in other disciplines to support his approach. In his words (2014, 349),

> A founding insight of both cognitive science and modern linguistics is that all human languages require supra-regular computational resources (resources above the finite-state level) ([Chomsky 1956, 1957]). This means that although a finite-state automaton can solve many useful problems (e.g., learn a lexicon, recognize word strings, etc.) there is a substantial class of problems that it cannot solve. These include all aspects of language in which flexible, extendable trees are needed as data structures, or where tree-identification and processing are core computational problems. For systems which rely strongly upon flexible, nested hierarchical structure, as do language and music, such tree-based processing is indispensable.

Thus far, my discussion of Fitch and Hauser's AGL experiment shows an novel application of FLT (i.e., to detect supra-regularity) with similar theoretical commitments from their predecessors. I have also discussed (1) the perceived advantage of applying a formal tool (i.e. FLT) in the study of the evolution of language and (2) the origin of a rising interdisciplinary research program in cognitive biology based on a new version of the AGL experimental protocols. In the next section, I argue that Fitch and Hauser's application of FLT can be viewed as tool-adaptation.

5.3 A Tool-adaptation

5.3.1 An Augmented Usage Profile

The novelty of Fitch and Hauser's application of FLT is evident. Indeed, I argue that the duo's 'experimental use' of the Chomsky hierarchy together with the AGL experimental protocols has expanded the usage profile of FLT. Specially, while formal grammars were present in Millier's Project Grammarama, Fitch and Hauser were the first to incorporate the idea of 'supra-regularity' from Chomsky's work to the design of their experiment. For this reason, I refer to Fitch and Hauser's version of the experimental protocols as the AGL powered by FLT or 'FLT-powered-AGL.' Moreover, with this new experimental use of FLT, I recast the instrumental character in four terms as opposed to three. That is, in addition to 'problem,' 'analytic result,' and 'solution,' there is a new step, which I call the 'experimental result'. To illustrate, in Fitch and Hauser's application of FLT:

- the 'problem' is concerned with the evolution of the ability to process hierarchical structures, such as 'sentences,' 'phrases,' 'words,' and so on.
- the 'analytic result' is that the stringsets generated by A^nB^n and $(AB)^n$, respectively, provide a good pair of tool for detecting supra-regularity.
- the 'experimental result' from the FLT-powered-AGT shows that tamarins failed to spontaneously recognize the A^nB^n grammar.
- the 'solution' is an interpretation of the experimental result coupled with the analytic result, i.e., the tamarins does not have supra-regularity, which means they do not possess the ability to process hierarchical structures.

This addition step in my illustration shows that FLT has given a new use in Fitch and

Hauser's (2004) application.

Another novelty resulting from the protocols of the FLT-powered-AGL is what I call a 'conceptual crossover.' For those who are familiar with the history of psychology, especially Chomsky's (1959a) criticism against the behaviorist Skinner's methodology, this novelty is the joining forces of formal models from generative linguistics and nonhuman animals as model organisms from the behaviorist tradition to the study of the evolution of human language. In the next section, I will argue that what seems to license such a conceptual crossover is a similar target profile in Fitch and Hauser's application of FLT.

5.3.2 A Similar Target Profile: 'It's All About Information Processing'

Like their predecessors, Fitch and Hauser hold that information processing is both computational and physical in nature. The brain, be it biological or silicon, is an information-processing device, and the mind is the program that runs on the device. Moreover, information-processing systems can be meaningfully classified into a containment hierarchy—meaningful because the classification is based on the correlation between the generative power of the grammar and the capacity of the recognizer as shown in the Chomsky hierarchy. Furthermore, there is a direct, information-processing impact on the recognition of patterns in acoustic sequences, such as parsing. (See Section 4.2.3 for my discussion on the 'correspondence hypothesis.') Fitch and Hauser's interpretation of the experimental results entails that to them, such an impact, while direct, manifests as the abilities of the tamarin monkey's 'internal program.' To illustrate, consider the two diagrams in Fig. 5.2.

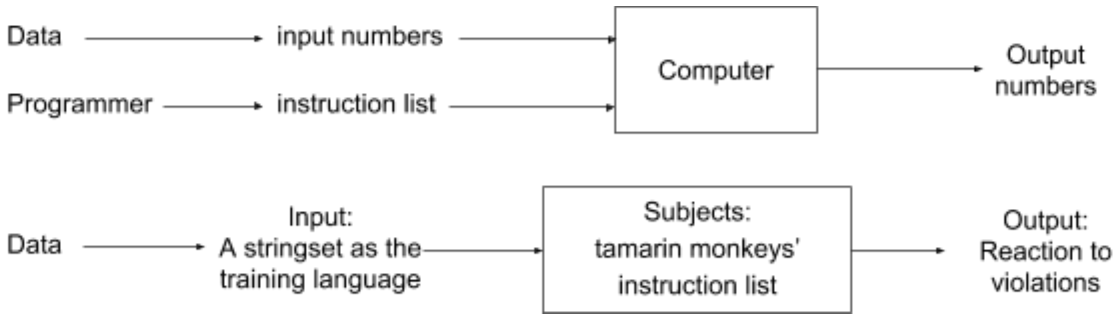


Fig. 5.2 Two Diagrams of Data Flow

In Fig. 5.2, I compare two different approaches to study information processing, one in computer science, another in the strain of cognitive biology discussed thus far. The diagram on the top is a replication of Ledley's (1962, 7) diagram captioned "Data flow to computer" (Section 4.3.1). The diagram at the bottom is my representation of Fitch and Hauser's version of the AGL experimental protocols. In particular, subjects are represented as the 'computer,' the training stringset as the 'input numbers,' and subjects' reaction toward violations as the 'output numbers.'

One thing to notice from this comparison is the internalization of the 'instruction list.' In computer science (the diagram on top), the instruction list is considered as input, whereas in cognitive biology (the diagram at the bottom), this list is 'internalized' in the subjects. In Fitch and Hauser's interpretation, it is this 'internal factor' that explains the output behavior observed in the experiment.

Moreover, to illustrate in what sense the AGL experiment protocols reveal the 'internal constraint' to supra-regularity, consider the diagrams in Fig. 5.3.

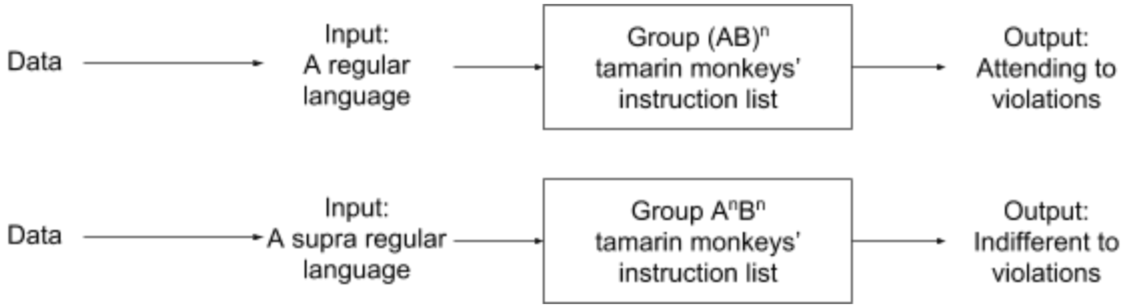


Fig. 5.3 A Comparison between Group $(AB)^n$ and Group A^nB^n

In Fig. 5.3, the diagram on top indicates that monkeys trained with a regular language (i.e., Group $(AB)^n$) are attentive to violations. In contrast, the diagram at the bottom shows that monkeys trained with a supra-regular language (i.e., Group A^nB^n) failed to show interest to violations. This contrast in monkeys 'output' behavior is interpreted as a limitation of their internal program. That is, their internal program could not recognize the supra regular training language. As the authors argue, the only difference between these two 'inputs' is the grammar that generates them, one being regular another supra-regular. Thus, to Fitch and Hauser, the diverging behaviors as output have but one explanation — tamarin monkeys' internal program is *not* supra-regular. In other words, to explicate their reasoning, if a system can solve a regular task but fail at a supra regular one, despite the ingredients of these two tasks are otherwise identical, then the said system cannot be supra-regular.

While this 'supra-regular' reasoning seems to be a weaker form of the correspondence hypothesis, it certainly shows a trace of influence from computer science as opposed from psycholinguistics. As Fitch and Hauser (2004, 378) put explicitly, "the AnBn grammar requires additional computational machinery beyond a finite-state automaton. In computer science terminology, this addition would minimally be a push-down stack." In other words, the computation constraints from which the tamarin monkeys suffer is, according to Fitch and Hauser, such a stack is absent both in their internal program and in their brain. Note that according to their pilot study, human subject have no difficulty of recognizing the violations. By this contrast, humans have the liberty of enjoying the push-down stack internally programmed and implemented in the brain for our disposal (Fitch 2014).[45]

To summarize, Fitch and Hauser's AGL experiment has created a new, experimental use of FLT, but their application has not changed the definitions of the components in FLT. The tool and its components remains to be units of information and rules of how smaller units compose bigger units. These components allow users to think about information processes, be it psychological, neurological, or implemented by silicon. Thus, Fitch and Hauser's application of FLT results in an augmented, and hence, a different usage profile, while the target profile stays largely the same. For this reason, it is suitable to view the case of FLT in cognitive biology as another example of tool-adaptation.

---

[45] There are other reasons why humans may not have a linear-bounded storage implemented in the brain.

Chapter 6

Epistemic Risks: Resistance and Backlashes from the Intended Users

In the previous chapter, I focused on one positive impact of tool migration. My analysis suggests that the combination of a similar target profile and an augmented usage profile enables the 'conceptual crossover' that brought to life an interdisciplinary research program — where animal model organisms meet formal models from generative linguistics in the study of the evolution of language. In this chapter, I discuss two negative repercussions: resistance and backlashes from the intended users.

Thus far, unlike Morgan's approach to traveling facts, my analysis has not explicitly addressed the aspect of users in tool migration, but the concern has been there: Both the inventors and the subsequent users are key figures to understand the tools in migration. Specifically, it is their respective applications of a given tool that give what I identify as the tool's target profile and usage profile. As I commented in Section 3.5, what appears to be a given tool borrowed and applied across disciplines turns out to be a wonderfully convoluted natural history, including the development and evolution of the tool, with possibly ever-changing profiles from one application to another. The story of game theory was one example. My investigation of formal language theory ('FLT') reveals the same trend, up until this point.

In this chapter, I discuss two reactions of Fitch and Hauser's work from their intended users: the neurolinguists and the experimental comparative psychologists. In Section 6.1, to begin, I briefly discuss what Fitch thought to be the appeal of applying FLT in the study of the evolution of language. In Section 6.2, as an example of resistance from intended users, I discuss a criticism of the stack model from alert neurolinguists. In Section 6.3, drawing form the migration story of FLT, I reflect on the relation between the autonomy and the limits of a research tool more generally. In Section 6.4, I come back to continue the story by reviewing the unfortunate search for 'recursion.' In particular, I argue that this search for 'recursion' (i.e., the capacity described as the 'infinite use of finite means' or 'open-endedness') constitutes a misuse of the experimental protocols of artificial grammar learning ('AGL') reported in Fitch and Hauser (2004). The misuse brought backlashes to Fitch and Hauser's work and cast a shadow on their achievement. Also in this section, I discuss Fitch's assessment of the source that has caused this misuse: a conflation of the concepts between 'recursion' and 'supra regularity.' While Fitch contends that this conflation is due to the mixed origins of FLT (including mathematical logic, linguistics, and computer science), I argue that this conflation shows up also in a paper he co-authored (Hauser, Chomsky, and Fitch 2002). Finally, in Section 6.5, I summarize the migration story of FLT. Drawing from these repercussions of FLT within cognitive biology that I reviewed in this chapter, I highlight the responsibility of users who are in charge of *importing* research tools. Being an importer of a research tool in migration is of a great responsibility. By doing so, one exposes his or her intended users to an assortment of epistemic risks, such as misusing the FLT-powered AGL experiment protocols to test for 'recursion.'

6.1 'Widely Accepted and Well-understood,' by Whom?

Fitch argues that FLT is "widely accepted, and well-understood" for classifying "computational systems ... including, presumably, actual brains ..." (Fitch 2010, 109).[46] He contends (ibid, 110) that the Chomsky hierarchy "provides a sensible starting point for classifying the computational power of different species in a way relevant to human language." In other words, to use my terminology, FLT has the appropriate target profile to begin with, and the AGL protocols add the 'experimental use' to its usage profile, both of which are indispensable to re-situate the tool in the study of the evolution of language.

Indeed, as I discussed in Section 5.3.2, FLT comes from the scientific tradition that approaches the mind as information processing, and the brain is generally assumed to be one type of physical devices that compute information. Thus, FLT can be said to have the *right* target profile — for the 'brain-as-a-computing-device' has been a theoretical commitment shared between the cognitive sciences and computer science. Similarly, the AGL protocols for testing supra-regularity can be said to have given FLT the *right* usage profile — for the Chomsky hierarchy has been a classification scheme used in both linguistics and computer science. The hierarchy has indeed been used to rank different such devices based on their generative powers. In other words, Fitch does have good reasons to think that FLT, especially the Chomsky hierarchy, is the right tool for the task, from the perspective of the users in linguistics and computer science, i.e., the disciplines where the tool was established. However, as I shown in Sections 6.2-3, FLT turns out to

---

[46] In (2010, 109), he argues that "the utility of [the] Chomsky hierarchy is that it provides a formally defined, widely accepted, and well-understood classification for computational systems less powerful than Turing machines (including, presumably, actual brains)."

be not so “widely accepted, and well-understood” by Fitch’s *intended users*. This result should not be surprising. In an interdisciplinary research program such as cognitive biology, a tool that might be well-understood, well-accepted by one group of users could very well be poorly-understood or even entirely rejected by another group of users. Let me begin by the story of the ‘IFG-as-stack model.’

6.2 Potential Limitations of FLT to Inform Neurolinguistics

6.2.1 Discrete Computation for the Brain?

The ‘IFG-as-stack model’ is explicitly formulated in Fitch (2014) but the idea appears earlier in Fitch and Friederici (2012). In both articles, Fitch justifies the model with the unique capacity in humans to process hierarchically ordered linguistic structure. Such a process is diagrammatically represented as the ‘derivation trees’ in both linguistics and computer science. Thus, the search for the implementation of the pushdown stack in the human brain suggests that the ‘derivation trees’ also have migrated with FLT to cognitive biology. However, in linguistics and computer science, derivation trees are used as means of analysis (e.g., explaining semantic ambiguity). To Fitch and fellow cognitive biologists, these derivation trees are real, not just psychologically and computationally, but also physiologically.

From this metaphysical commitment, a problem arises: the tree-like processes is a result of discrete modeling in generative linguistic, whereas the brain a continuous device. For this reason, some scholars worry that the Chomsky hierarchy might not be the right tool for studying the brain. For instance, consider Honing and Zuidema’s (2014, 376) critical response to Fitch:

> We do not share ... Fitch's faith in ... the branch of computing theory that takes trees (and the Chomsky Hierarchy) as central notions. ... [A]s we zoom in on the neural basis of this [tree] structure, important questions arise about the cognitive reality of the building blocks of tree-based descriptions: the symbolic nodes, the ordering of the branches, and the implied hierarchical levels.

The problem lies in the attempt to model the brain, a continuous system, with a discrete model that did not take into account the kind of 'stuff' that the brain is made of. Honing and Zuidema continue (ibid., 376):

> [U]ltimately, computing in the brain is based on electrical and chemical substrates that vary on continuous scales. How discrete and tree-like structures may emerge from such a continuous basis is an important research question for cognitive science … With exact answers to that question still lacking, ... it is important to realize that a continuous system may approximate the behavior of a discrete idealization to an arbitrary degree and still remains at heart a continuous system.

They (ibid., 376) conclude that the derivation trees:

> might thus provide a good description for some aspects of the behavior of a system, but fail for other aspects, and completely disappear when zooming in, because the primitive operations of the system are very unlike trees.

In other words, before taking the IFG-as-stack model seriously, one might wish to furnish it with an account of how the discrete structures emerge from the continuous processes in the brain.

6.2.2 A Stack Model for Linguistic Memory in the Brain?

Trying to locate the pushdown stack in the brain assumes that the human brain implements linguistic memory with a mechanism like the last-in-first-out stack (see Section 4.1.4). Essentially a product of engineering, a pushdown stack is an abstract data type, invented for automatic processing of coded data, and the computing machine for performing the method.[47] Similar to the working tape of a Turing machine, a pushdown stack is a linear storage that behaves in a particular way: By design, it pushes items down the stack and pops the latest one out before an earlier item may be retrieved.

As the psychologist Jonathon Crystal (in personal conversation) points out, for a stack to be a plausible model of human linguistic memory, one would expect the last-in-first-out fashion of memory recall in humans. That is, humans would perform better recalling items of later part of a sequence than recalling items of the beginning of the sequence. However, a phenomenon called the serial-position effect studied by psychologists counters such an expectation.

Serial-position effect refers to an observation about memory recall. It has been shown that when people hear a reasonably long list of unrelated words and are asked immediately to recall them in *any* order, they tend to perform similarly well recalling the words at the end *and* at the beginning of the list. The middle of the list is “far less likely to be recalled” (Hardy and Heyes 1999, 63). Thus, it remains questionable whether the concept of a stack alone is appropriate for understanding the neural storage of linguistic

---

[47] In 1957, Friedrich Ludwig Bauer and Klaus Samelson filed for a patent of the “Method for the automatic processing of coded data, and computing machine for performing the method.” https://patents.google.com/patent/DE1094019B/en

information with regard to production and comprehension.

6.2.3 The Place of Infinite Processing Time in Cognitive Neuroscience

Infinity is a crucial modeling choice to the theoretical development of FLT. When Alan Turing (1936) models computation with the concept of a machine, both its memory and the time available for the machine to complete the computation are assumed to be infinite (Section 4.1.4). The working tape of a Turing machine is infinite in length. And the computability of a function is defined based on whether "there exists a set of instructions that will result in a Turing machine computing the function regardless of the amount of time it takes" (Barker-Plummer 2012). Chomsky (1956, 115) has also explicitly stated that "[i]n general, the assumption that languages are infinite is made for the purpose of simplifying the description" i.e., the grammar. For the "grammar must reflect and explain the ability of a speaker to produce and understand new sentences which may be much longer than any he has previously heard," he says in an endnote (Chomsky 1956, 124). However, while restricting memory becomes the key to classifying different classes of automata, the Chomsky hierarchy does not concern the factor of computing time. Indeed, in computer science, the Chomsky hierarchy is entirely concerned with 'computability,' whereas 'computational complexity,' which deals with the aspect of processing time (among other things), is considered in a separate context.

On the other hand, regardless of species, biological brains compute with limited time. It thus poses a question whether FLT indeed provides the right tool for zooming in to the neural activities in addition to merely classifying them based on their performance. As Honing and Zuidema suggest (Section 5.2.1), how the brain computes and stores information remains itself a research question. Relying on the ‘IFG-as-stack’ model or FLT to inform research in neurolinguistics could be placing the cart before the horse.

6.3 The Undetermined Autonomy and Limits of the Tools: A Discussion

6.3.1 Modeling Choices Carried Over to Become Theoretical Commitments

Discussed in Section 6.2, those theoretical commitments concerning discrete computation and infinite computing time were modeling choices made by inventors of the tool in the disciplines where FLT was initially established and further developed. In a sense, mathematical logic can be viewed as FLT’s ‘native’ discipline whereas linguistics and computer science its ‘home’ disciplines.

Having migrated to cognitive biology, those modeling choices became theoretical commitments. Together with the AGL experimental results, it gives rise to the IFG-as-stack model for exploring how the brain implements a stack, which in turn motivates the search for a biological explanation for human linguistic capacity. My study of FLT confirms one thing: Research tools develop and evolve over the course of being used, within or outside their native disciplines. Such a discovery strengthens the main point I argued in Chapter 3: Keeping the integrity of a migrating tool is not necessarily desirable. Focusing on the changes to a tool (especially changes to either the target profile or the usage profile, or both) enriches our understanding of tool-use in science.

6.3.2 Limits of a Tool

That said, I should note that no users (not even the inventors) have the *total* authority over what a tool *could* or *could not* be used to do. For instance, Maynard-Smith’s criticism about economists using evolutionary game theory (‘EGT’) is an example of misjudgment — it was an inventor underestimating the potential of his creation, as I argued in Section 3.4.5. Unbeknownst to him, certain modification to EGT’s target profile (allowing, for example, for fitness without biological heredity) has made justifying its applications in social sciences possible. However, there may be cases in which the accumulated modifications thus far still would not justify using a tool in a certain way. If true, it follows that there may be a limit to each tool’s functionality (i.e., to what it could or could not do), with or without modifications. To illustrate, consider a simple example. An adjustable wrench is a versatile tool. It can spin bolts of varying sizes. In addition, it could function as a clamp. With its weight, it could be used as a hammer, and with its length, it could function as a unit of measurement if necessary, and so on. However, this list of its possible uses is likely to end somewhere. For instance, without an overhaul, a wrench could not be used as a roller because it does not have the shape for such a use. And there are some things that it simply could not be, such as a Mars space-probe. Of course, finding the limit of a tool is not straightforward (which is why I did it arbitrarily in the example of a wrench). Indeed, when modifications to a given tool are allowed, it is not clear whether there will be a fixed limit of the tool’s functionality. Thus, between the inventor’s intended use(s) and the tool’s limit that may or may not be fixed, there is plenty of room for the users’ creativity.

6.3.3 A Misuse of a Tool Due to a Carried-over Conflation of Concepts

Still, some misuses of a tool are relatively obvious to pick out. In the next section, I discuss an example in which the users misjudge the capacity of the tool. This example also shows that after all, FLT is not well-understood for classifying computing systems without qualification. It may be true that FLT is well-understood among theoretical computer scientists for classifying formal languages and their recognizers. However, in an interdisciplinary research program, it is dangerous to assume that such a tool will be similarly well-understood by another group of users who are new to it. The AGL experimental protocols powered by FLT are such a tool. After 2004, some researchers started using it mistakenly to show that 'recursion' — a computational feature that Fitch, Chomsky and Hauser hypothesize (2002) to be unique to humans and unique to language — exists in some nonhuman animals, when in fact, the tool does not have the capacity to detect recursion. I tell a story of this episode of tool-misuse in the sections that follow.

To get slightly ahead of myself, Fitch (2012) argues that the unfortunate search for recursion with the FLT-powered AGL experiment is misled by a pervasive conflation of key concepts in FLT, such as the Chomsky hierarchy and recursion. Fitch (2010, 84) suggests that this conflation appears in "the merger of mathematics and linguistics called formal language theory" (which I discuss in Section 6.4.5).

Without contradicting Fitch's statement, I argue that there was another source of confusion. Those researchers, who *mis*applied the experimental protocols for detecting recursion, were misled by a paper co-authored by Hauser, Chomsky, and Fitch (2002) (as I argue in Section 6.4.6). In other words, this conflation of concepts has been carried over to cognitive biology, which then starts the unfortunate confusion. The story begins with why, after 2004, there was suddenly a surge of testing recursion in nonhuman animals.[48]

---

[48] E.g., Abe et al 2011; Bahlmann et al 2008; de Vries et al 2008; Gentner et al 2006;

6.4 The Unfortunate Search for 'Recursion'

6.4.1 The Faculty of Language

It all started with the question: "The Faculty Of Language: What Is It, Who Has It, And How Did It Evolve?" (title of Hauser, Chomsky, and Fitch 2002). The term 'faculty of language' refers to a set of abilities that enable humans to acquire and use natural languages. Humans are thought to be endowed with these abilities based on our neurobiological and genetic composition. Traditionally, the faculty of language was conceived to be uniquely human. It is sometimes compared with 'animal instinct' (e.g., Pinker 1994). That is, using language is uniquely human in the same way that spinning webs is uniquely spider.

In contrast, Hauser, Fitch, and Chomsky (2002) propose a three-system concept of the language faculty. They argue that some or all of these systems may have an independent evolutionary history, recruited to bringing about linguistic abilities only in humans. Thus, studying animal model organisms may shed light on the evolution(s) of the faculty of human language. They begin with laying out the components of this faculty, all of which are meant to account for what language really is.

6.4.2 Sensory-motor System, Conceptual-intentional System, And Abstract Linguistic Computational System

The first component is the 'sensory-motor' system, responsible for hearing (or viewing) and speaking (or gesturing). The second is the 'conceptual-intentional' system, in charge

---

Hochmann et al 2008; Marcus 2006; Perruchet and Rey 2005; Rey et al 2012; Stobbe et al 2012; Udden et al 2012; van Heijiningen et al 2009.

of meaning. The last is the "abstract linguistic computational system," for creating internal representations of discrete expressions (ibid., 1571). Each of the three systems is responsible for one aspect of language, which according to the authors is "fundamentally, a system of sound-meaning connections" with "a finite set of elements" and "a potentially infinite array of discrete expressions" (ibid). In particular (ibid),

> Each of these discrete expressions is ... passed to the sensory-motor and conceptual-intentional systems, which process and elaborate this information in the use of language. Each expression is, in this sense, a pairing of sound and meaning.

One important step in their analysis of the faculty of language is the introduction of 'recursion.' "At the minimum," the abstract linguistic computational system "includes the capacity of recursion" for creating "an open-ended and limitless system of communication" (ibid., 1571/1578). They call this system 'the faculty of language in the narrow sense' or the 'FLN.' Thus, recursion is the main characteristic of the FLN. In contrast, the whole three-part system (i.e., FLN together with the sensory-motor system and the conceptual-intentional system) is called 'the faculty of language in the broad sense' or the 'FLB.'

6.4.3 The Evolution of the Faculty of Language and the 'Recursion-only' Hypothesis

Hauser, Chomsky, and Fitch (2002) argue that this FLB/FLN distinction helps to break down questions in the literature that concern the comparative and evolutionary aspects of the faculty of language. According to them, these questions were:

1. whether the faculty of language is uniquely human or shared with other species,

2. whether the evolution of such a faculty is gradual or saltational, and
3. whether it evolved as a unique adaptation for communication or for some other computational problems.

They suggest, for instance, that researchers may now investigate details of *each of the three subsystems* along the comparative aspect of the subject matter (as shown in Question 1). Those details may inform the study of the evolutionary aspects of the topic (as in Questions 2-3). That way, scholars may piece together a fuller understanding of the evolution of language.

With those traditional research questions renewed, the authors then offer hypotheses for further research. I re-state two of their proposed hypotheses as follows: Either FLB as a whole is a "uniquely human adaptation for language" (ibid., 1572), or "[o]nly FLN is uniquely human" (ibid., 1573), surely not both. The word 'only' in the latter makes the two hypotheses mutually exclusive. This latter hypothesis was interpreted as the 'recursion-only' hypothesis. As the authors put in the abstract (ibid., 1596): "We hypothesize [both] that FLN only includes recursion and [that FLN] is the only uniquely human component of the faculty of language."

6.4.4 Empirical Methods for Detecting Recursion in Nonhuman Animals

Verifying the recursion-only hypothesis is not straightforward. FLN supposedly contains only recursion, and recursion is supposedly responsible for an open-ended and limitless system of communication. To illustrate, scientists assume that humans have recursion. A typical argument goes as follows. For any given sentence, e.g., 'the Earth is round', one may create indefinitely many sentences by repeatedly adding a clause in front of it, such as ‘Mary believes that…,’ ‘John knows that…,’ ‘it is possible that…,’ and so on. It appears that a test for recursion in a subject requires a communication system between the experimenter and the subject. Thus, not speaking any animal languages, and indeed not knowing whether any exist, scientists wishing to detect recursion in nonhuman animals need to resort to other signs of recursion.

Hauser, Chomsky, and Fitch (2002) suggest two possible directions: number representation and rule learning. Because both abilities “can be investigated independently of communication,” experiments of these two abilities can “provide hints as to the nature of the constraints on FLN” (2002, 1577). This statement is one of the first places where confusion starts to loom.

Open-endedness in Number Representation

According to the authors, the best evidence for number representation in nonhuman animals at the time was from work done by Boysen and Matsuzawa on chimpanzees. (They cite Boysen and Bernston 1989, Kawai and Matsuzawa 2000, Matsuzawa 1985). However, unlike human children, Chimpanzees do not show the open-endedness in their number representation. The authors continue (1577),

> A human child who has acquired the number 1, 2, and 3 (and sometimes 4) goes on to acquire all the others; he or she grasps the idea that the integer list is constructed on the basis of the successor function. For the chimpanzees, in contrast, each number on the integer list required the same amount of time to learn. In essence, although the chimpanzees’ understanding of Arabic numerals is impressive, it parallels their understanding of other symbols and their referential

> properties: The system apparently never takes on the open-ended generative property of human language.

What might be the cause of such a fundamental difference? The authors consider two possibilities. The first possibility concerns the difference in learning experience due to external factors, a view that they attribute to Carey (2001). For instance, children "typically learn an arbitrar[il]y ordered list of symbols ('1, 2, 3, 4 …'), and only later do they learn the "precise meaning of such words" (Hauser, Chomsky, and Fitch 2002, 1577.). In contrast, "apes and parrots were taught" by the experimenters "the meanings one by one without learning the list (ibid., 1577). Thus, this difference from the training regime of animals might have resulted in their fundamentally different experience.

The second possibility concerns the difference in the computational constraints of the nonhuman animals, i.e., "the kind of statistical inference that animals can compute" (ibid., 1577). Importantly, this is where the Chomsky hierarchy, rule learning, and the thesis of supra-regularity (Sections 4.2.4 and 5.2.3) enter the discussion, all (wrongly) in connection with recursion. After introducing these concepts (which I quote below), the authors state in a caption that "in parallel with the faculty of language, our capacity for number relies on a recursive computation" (ibid., 1576). The passage that may have conflated the concepts of supra regularity (or 'hierarchical processing' in their terminology) and recursion reads as follows (1577, emphasis mine):

> Early work in computational linguistics (Chomsky 1975, 1956, Chomsky and Miller 1958) suggested that we can profitably think about language as a system of rules placed within a hierarchy of increasing complexity. At the lowest level of

> the hierarchy are rule systems that are limited to local dependencies, a subcategory of so-called "finite-state grammars." Despite their attractive simplicity, such rules systems are inadequate to capture any human language. Natural languages go beyond purely local structure by including a capacity for *recursive embedding of phrases within phrases*, which can lead to statistical regularities that are separated by an arbitrary number of words or phrases. Such long-distance, hierarchical relationships are found in all natural languages for which, at a minimum, a "phrase-structure grammar" is necessary. It is a foundational observation of modern generative linguistics that, to capture a natural language, a grammar must include such capacities (authors pointing to a figure which I replicate in Fig. 6.1).

This passage is misleading for two reasons. First, it seems to suggest that supra regularity, which the finite-state grammars lack, distinguishes between formal systems that have open-endedness and formal systems that lack open-endedness.[49] This suggestion is incorrect because both finite-state grammar and phrase-structure grammar are open-ended; in fact, all four classes of formal grammars in the Chomsky hierarchy are. After all, as Fitch states later (2014, 347): FLT is "the study of infinite sets generated by finite means." Thus, these two classes in the Chomsky hierarchy cannot inform scholars about recursion in terms of open-endedness. Second, it invokes the term 'recursive' referring to recursively "*embedding ... phrases within phrases*," whereas by 'recursion' they conflated the notions by thinking of open-endedness as due to recursion

---

[49] According to the three authors (2002), recursion is supposedly about the capacity of handling an open-ended system, which could be integers, numbers in general, or natural languages.

(Hauser, Chomsky, and Fitch 2002, 1577, emphasis mine). Let me elaborate on this in the next section.

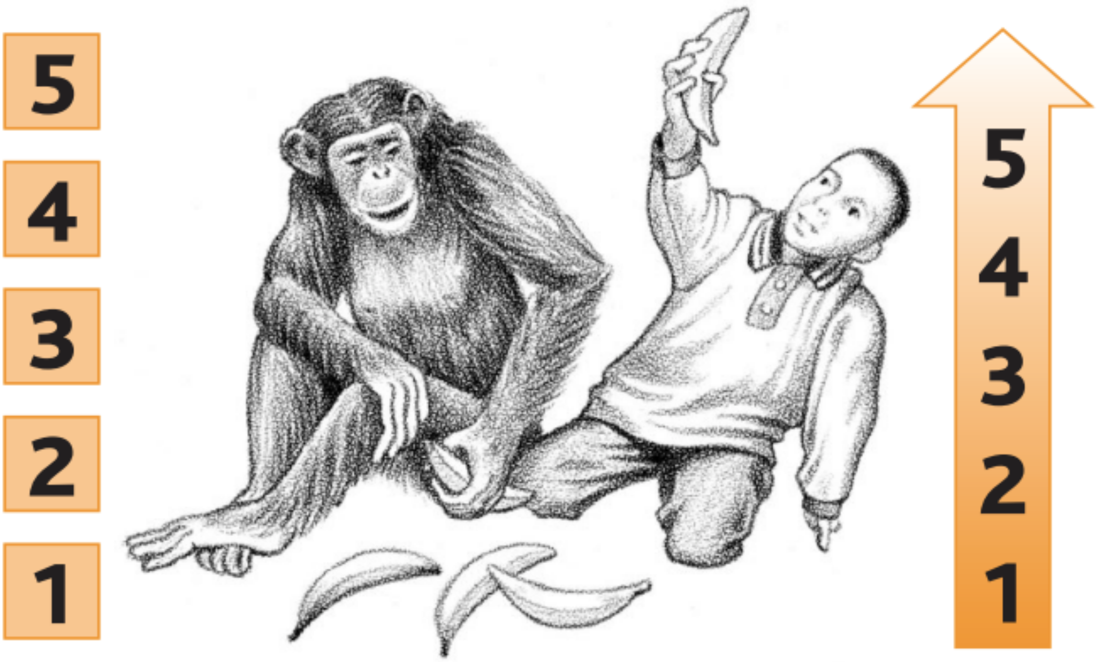


Fig. 5. Human and nonhuman animals exhibit the capacity to compute numerosities, including small precise number quantification and large approximate number estimation. Humans may be unique, however, in the ability to show open-ended, precise quantificational skills with large numbers, including the integer count list. In parallel with the faculty of language, our capacity for number relies on a recursive computation. [Illustration: John Yanson]

Fig. 6.1 Recursion as Open-endedness. This illustration is captured from the Fig. 5 in Hauser, Chomsky, and Fitch (2002, 1576).

6.4.5 Recursion: Open-endedness and Self-embedding

Within their paper (Hauser, Chomsky, and Fitch 2002), recursion has at least two readings: recursion as the capacity for using an open-ended system and recursion as self-embedding. The latter reading implies the former, but not vice versa. In linguistics, 'recursion' refers to the property of a rewriting rule being 'self-embedding,' which is indeed a supra regular feature (Section 5.2.3). However, 'self-embedding' could describe both stringsets A^nB^n and $(AB)^n$ in Fitch and Hauser's (2004) AGL learning experiment. In particular, a rewriting rule is 'self-embedded' if the same phrase type, *S*, appears on both sides of a production in the form $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$. As such, the stringset generated by $(AB)^n$ *can* be described by a recursive rule. To illustrate, consider the three productions shown in (5.1).

(5.1)

i. $\boldsymbol{S} \rightarrow \mathrm{A}\,\boldsymbol{S}\,\mathrm{B}$

ii. $S \rightarrow \mathrm{AB}$

iii. $\boldsymbol{S} \rightarrow \boldsymbol{S}\mathrm{AB}$

In both (5.1i) and (5.1iii), the phrase ***S*** appears on both sides of the production arrow, making the two of them self-embedded rules, whereas (5.1ii) lacks such a pattern.

Note that a grammar consisting of rules (5.1i) and (5.1ii) can generate the stringset of A^nB^n, whereas a grammar containing both (5.1ii) and (5.1iii) can generate the stringset $(AB)^n$. As such, both stringsets are said to be generated by a self-embedded rule. Recursion in terms of self-embedding does not tell apart the stringsets in Fitch and Hauser's AGL experiment.[50] Thus, while the FLT-powered-AGL experimental protocols developed by Fitch and Hauser may have the right profiles for detecting supra-regularity, the tool certainly does not detect recursion in terms of open-endedness (because all four classes of grammars in the Chomsky hierarchy are open-ended), nor does it detect self-embedding (Fitch 2010). Crucially, while supra-regularity is irrelevant with regard to 'recursion as open-endedness,' it does concern 'recursion-as-self-embedding.' A

---

[50] What they should have said is 'center-embedding' instead of 'self-embedding'. However, recursion in linguistics is not limited to 'center-embedding'. Thus, either way, confusion persists.

grammar consisting at least one self-embedding rule disqualifies as Type 3. Thus, at the end of discussing number representation, ‘recursion’ gains an additional reading of being about self-embedding, a reading that is used in linguistics for describing rewriting rules.

6.4.6 Misinterpretation of the Experimental Results

Recall my analysis of FLT-powered-AGL experimental protocols (Section 5.3.1). The experimental result would reveal whether the subjects spontaneously recognize their training grammar. When the subjects do not recognize a supra-regular grammar, everything else being equal, this result is interpreted as that the subjects do not possess the addition computational resource to be supra regular Due to the discussion in Hauser, Chomsky, and Fitch (2002) that potentially conflated supra-regularity and ‘recursion as open-endedness,’ commentators took Fitch and Hauser’s (2004) work to be confirming the recursion-only hypothesis (Section 6.4.3). Fitch reports several such cases (2010, 87):

> The first place I saw this misinterpretation was ... written by David Premack, appearing in the same issue of *Science*: “In ... this issue, Fitch and Hauser ... report that tamarin monkeys are not capable of recursion. Although the monkeys learned a nonrecursive grammar, they failed to learn a grammar that is recursive. Humans readily learn both” (Premack 2004, 318). ... [A] latter paper by Perruchet and Rey (2005) apparently assumed that we shared Premack’s assessment of our experiment, disputing our “claim” about A^nB^n “generating center-embedded sentences.” Most recently, Gentner et al. (2006) concluded that starlings are able, with training, to master the A^nB^n grammar, and titled their paper “Recursive

> syntactic pattern learning by songbirds." The abstract states that A^nB^n is a recursive, self-embedding, context-free grammar' " The accompanying commentary by Gary Marcus continues the error, asserting that "The A^nB^n language … is generally assumed to be recursive …."

To Fitch, the thought that A^nB^n requires recursion not only is incorrect, but also "appears to reflect an inadequate grasp of computer science and formal language" (2010, 87). Most importantly, he says, "our paper [2004] did not mention recursion, because A^nB^n is not a test for recursion" (ibid.). He continues, "the question of recursive implementation is orthogonal to the analysis of grammatical power embodied in the Chomsky hierarchy" (ibid.). What A^nB^n grammar does in the AGL experiment is to test "the additional memory mechanism(s) to keep track of 'n' ", and "it is an error to see success at recognizing strings from A^nB^n (as for starlings) as indicating recursion, or to see failure at the same task (as for tamarins) as necessarily implying a lack of recursion" (ibid., 87-88).

While being polemic in his response, Fitch does "regret certain editorial decisions," not in the (2004) but in the earlier paper (Hauser, Fitch, and Chomsky 2002) that made their "argument more opaque than desirable" (2010, 75). In the next section, I discuss the key passages (in Hauser, Chomsky, Fitch 2002) where their misleading message lies. He is correct to regret — it was him and his co-authors who exposed their scientific community to the risk of making the mistake. (And let's not forget that one of his co-authors is the inventor of the initial hierarchy!)

6.4.7 Conflation of Concepts Carried Over by the Importers

The most likely source that started *mis*applying A^nB^n grammar to be a test for recursion, I argue, appears in Hauser, Chomsky, Fitch (2002). Fitch is correct; Hauser and Fitch (2004) did not mention recursion. However, that paper was not the time the scientific community first heard of their novel AGL experiment. In 2002, Hauser, Chomsky, and Fitch disclosed the experiment and discussed their conclusion (then in preparation). At the end of the relevant passage (which I quote below in its entirety), they added two lines that are unmistakably misleading (ibid., 1578, omitting descriptions of the experimental procedure, emphasis mine):

> Fitch and Hauser [then in preparation] recently completed a study comparing finite-state and phrase-structure grammar acquisition in human adults and tamarins. ... The phrase-structure rule tested was AnBn. ... Results showed that human adults rapidly learned this rule implicitly, distinguishing consistent and inconsistent strings. Tamarins, in contrast, failed in three separate experiments testing their ability to acquire this grammar, but they readily mastered a finite-state variant (ABn)[51] implemented with the same stimuli and testing conditions. This suggests that tamarins have a limited capacity to learn the type of long-distance hierarchical dependencies necessary to achieve the class of phrase-structure grammars. If true, this limitation would place severe restrictions on their capacity to learn any natural human language. It is currently unclear whether this limitation generalizes to other animals, and whether it is similarly imposed on humans at different stages of development. *Nonetheless, such experiments provide an empirical approach to exploring key differences between humans and animals relevant to FLN.*

---

[51] Typo in the original; it should have been $(AB)^n$.

If FLN contains only recursion, and if the AGL experiments “provide an empirical approach to explore” important species differences “relevant to FLN,” then it seems natural to conclude that the AGL experiment was the go-to tool for testing recursion.

6.5 Summary of the Migration of the Theory of Formal Language

The tool of FLT-powered-AGL experimental protocols successfully promoted the comparative approach and interdisciplinary research in the study of language evolution. The tool even allows nonhuman animal model organisms to re-enter the study of human language, and vice versa. Hauser, Chomsky, and Fitch state that (2002, 1578):

> Our review has stressed the usefulness of animal data for theories about humans, but this exchange need not be one-way. As the research program we have sketched progresses, more general principles about cognitive evolution may emerge.

Interdisciplinary collaboration is the key to making progress in this endeavor. They urge scholars including “[l]inguistics and biologists, along with researchers in the relevant branches of psychology and anthropology” to “move beyond unproductive theoretical debate to a more collaborative, empirically focused and comparative research program” for the goal of “uncovering both shared ... and unique components of the faculty of language” (2000, 1578). Unfortunately, in articulating such a research program, the authors inadvertently introduced a conflation of concepts between supra regularity and recursion — a conflation which Fitch believes to have existed before their work — leading to yet another unproductive debate.

After much effort devoted to clarification (including an ERC-sponsored workshop and a number of publications), the rush to find recursion seems to have subsided. Looking back, Fitch (2014, 30) notes: Although my colleagues and I have previously discussed the generation of new [hierarchical] levels in terms of "recursion" [his citation to Hauser, Chomsky, Fitch 2002], I now regret this."

Recursion is leaving the spotlight, taken over by the stack model of the brain. The AGL experiments and FLT remain at the center of cognitive biology (e.g., Jiang et al. 2018, Fitch 2018). One lesson learned: unlike Fitch once thought (2010, 109), FLT was not necessarily "well-understood" for doing what he proposes to do. Indeed, within an interdisciplinary research program, whether a tool's is 'well-understood' cannot be estimated disregarding the users who are invited to use it.

Chapter 7 Conclusion

Tool Migration is Epistemically Risky — Having A Similar Target Profile Does Not Seem to Help

7.1 Summary

This dissertation focused on the scientific practice in which a mathematical construct that was originally developed to study a particular subject matter but subsequently used in other disciplines or subdisciplines for a different subject matter, a phenomenon that I call 'tool migration.' The main motivation behind this research was the concern that tool migration can be 'epistemically risky.' More specifically, that uprooting a research tool from one disciplinary context and re-situating it for use in another can change how the tool is applied; whatever has made the tool useful and reliable in the first place may not have stayed the same in the new context. The two migration stories illustrated in this dissertation legitimize this concern. The epistemic risks associated with tool migration are real. However, my analyses of these tool migration stories reveal that changes in the tools due to uprooting or re-situation are not necessarily the problem; instead, appropriate and proactive modifications to a tool in migration are conducive to successfully applying an old tool in a new context. By contrast, having a similar set of definitions concerning the theoretical terms in the formal construct (i.e., the tool's target profile) does not make tool migration less precarious.

The epistemic risks associated with tool migration have not been explicitly addressed by philosophers of science. The lack of attention to this topic is surprising in light of the recent and growing interest in topics such as scientific model transfer. For instance, the discussion of 'model transfer' concerns a relatively small set of mathematical models that are applied in multiple disciplinary contexts. Humphreys (2004, 2018) argues that models which are transferred to study phenomena of different domains owe their versatility to the computational tractability they afford. In contrast, Knuuttila and Loettger (2014, 2016) suggest that in addition to tractability, versatile models also offer conceptual frameworks for theorization, which they label 'model templates.' However, while 'model transfer' is similar to what I call tool migration, prior analyses have not dealt with the risks inherent in this aspect of scientific practice.

In a slightly different context, Morgan (2010) proposed a similar warning about the use of evidence outside its site of construction or discovery, or what she calls 'traveling facts.' Her argument that 'losing integrity' poses a threat to the usefulness and reliability of traveling facts coincides with my preconception about tools in migration. Her warning is based on a collection of case studies (Howlett and Morgan 2010) devoted to understanding the different ways in which facts are received and treated as evidence according to different disciplinary bases. It was clear to me that my reservation about the use of the tools in migration resembles Morgan's concern about the use of the traveling facts. I thus modeled my investigation to examine whether 'losing integrity' also poses a threat to the usefulness and reliability of the tools in migration.[52]

---

[52] What was not clear to me is how my concern of the 'importers' in tool migration turns

The main contributions of the current project are 1) the analytic framework that I developed and 2) the findings I discovered by applying the framework to my case studies. On the analytic framework, first I characterized an application of a research tool in three types of tasks: problem formulation, calculation (or computation), and interpretation. Specifically, in an application mathematical constructs are used to:

(i) formulate a research question into a tractable form,

(ii) obtain an analytic result of the formulation, and

(iii) interpret the analytic result as a solution or answer to the research question.

Second, within an application of a tool, I distinguished the tool's:

- Target profile: the set of theoretical terms (especially their definitions and the relations between these terms) that factor into formulating the problem, and
- Usage profile: the set of ways in which the results of calculation are interpreted.

Between applications, I distinguished:

- Established application, and
- Novel application.

Third, with these two concept-pairs, I build a four-fold typology to check for the 'integrity' of a tool in migration. Specifically, I compared the similarities or differences in either profiles between applications.

- When a novel application retains both the target profile and usage profile of a tool from the tool's established application(s), I classify it as tool-application,
- when the novel application retains a critically similar target profile of the tool but

---

out nicely corresponding with Morgan's concern of the 'curator' and 'labeling practice' in traveling facts. See footnote 8 for discussion.

has changed the tool’s usage profile has changed, I call it tool-adaptation,

- when the novel application retains a critically similar usage profile of the tool wut has changed the tool’s target profile, it is a tool-transfer, and
- when the novel application has changed both the target profile and usage profile of the tool, it is a tool-transformation.

I then applied this framework to argue that the migration stories of game theory and formal language theory are examples of tool-transformation and tool-adaptation, respectively.

7.1.1 Case Study 1: Game Theory

Game theory was constructed for studying strategic interaction between agents capable of making decisions (von Neumann 1928); yet game theory has since been used in evolutionary biology to study the change of phenotypes in a population of organisms that may or may not be capable of making decisions (Maynard-Smith and Price 1972).

From the migration stories of game theory, I found substantial changes, especially the changes in the definitions about the ‘players’ and the ‘payoffs.’ To use my terminology, between these applications, there is a collection of diverging target profiles. For instance, once game theory left the context of the social sciences and entered the context of evolutionary biology, the ‘players’ are no longer defined to be decision-making individuals, nor is the ‘change of strategy’ a result of individual player’s decision-making process. Such a definition (i.e., players being capable of choosing one strategy over another) simply does not apply to subject matter of evolutionary biology. Moreover, a ‘strategy’ in evolutionary game theory refers to a phenotype that individual organisms display, and the concept of a ‘payoff,’ in turn, is redefined as Darwinian fitness (i.e., the number of copies that it will leave to play in the games of a succeeding generation). Consequently, in evolutionary biology, the phrase ‘a change of strategy’ means a change of the proportion of the phenotypes in a population. Furthermore, in order to account for this change in the proportion of phenotypes, it is assumed that (Maynard-Smith and Price 1972) an individual organism passes on its phenotype to its offspring according to the fitness (i.e., payoff) assigned in the payoff matrix. This became known as the heritability assumption about the individuals in evolutionary game theory. As a result, the change in the proportion of phenotypes is understood as a result of differential biological reproduction instead of decision-making. For this reason, when the social scientists started applying evolutionary game theory to theoretical economics, Maynard-Smith (1997) criticized their applications for being ignored of the heritability assumption.

My analysis showed that Maynard-Smith’s criticism was a misjudgment because there has been a change in how ‘payoff’ is defined. Such a change in the definition of ‘payoff’ can be viewed as ‘counter-actors’ to offset other changes (e.g., the subject matter changes from being about decision-making individuals to being about phenotype-displaying organisms). Had it not been this counter-acting modification in the definition of ‘payoff,’ Maynard-Smith’s (1997) criticism would have been justified. Specifically, according to Maynard-Smith (1997), evolutionary game theory “hinges on the notion of heredity, i.e. the essential notion being that your success in the game determines how many children you have, and your children are like you, and the whole thing [i.e., the propagation of a particular strategy] hinges upon that essential assumption." In other words, if the social scientists used the tool to study human affairs in merely a 'plug-and-play' manner, they have mischaracterized their human subjects. However, unbeknownst to Maynard-Smith, Taylor and Jonker (1978) in their application contested the idea that the propagation of a strategy can be understood *only* as a result of heredity, as Maynard-Smith did. According to them, it can be understood *also* as a result of social learning (e.g., imitation). Thus, this change of assumption, which changes the target profile of the tool, in turn properly re-situates the tool from evolutionary biology back to the social sciences.

In general, counter-acting modifications to a tool in migration, despite causing the tool to 'lose integrity,' are helpful to prevent errors that are particularly associated with tool migration. In other words, counter-actors are crucial (and perhaps indispensable!) to properly re-situate a tool in a new context. As mentioned in Chapter 3, my finding of 'counter-acting modification' in tool migration is consistent with the existing literature on the improvement of 'theoretical templates' (Humphreys 2018) and the construction of model templates (Knuuttila and Loettger 2014, 2016).

7.1.2 Case Study 2: Formal Language Theory

From the migration stories of formal language theory ('FLT'), I found a similar set of definitions concerning the theoretical terms: 'alphabet,' 'words,' 'phrases,' 'sentences,' 'rules,' 'derivations,' etc. Between linguistics and computer science, these theoretical terms are units of information that are hierarchically organized. Specifically, in linguistics, an alphabet is a set of symbols, which are the smallest units that form words, which are units that form phrases, and phrases, in turn, are units that form sentences. In computer science, as I illustrated with the programming language ALGOL, characters are the smallest units that which form expressions, and expressions, in turn, are units that which form statements. In both disciplines, grammatical rules are descriptions of the allowable forms, and derivations are the processes that generate these forms. These similarities suggest that the target profile of FLT has survived through the migration from linguistics to computer science. A similar set of assumptions can be found in Fitch and Hauser's (2004) design of their experiment. They created two classes of syllables, which are then used to generate strings as experimental stimuli based on rules that they call 'grammars.'

However, in contrast to a converging target profile, I found increasingly different ways in which FLT is used across the disciplines. To use my terminology, between these applications, there is an augmented usage profile. My analysis suggests that in each of those disciplines, a 'new use' is added to the tool in their respective application. Specifically, there are the 'theorizing-use' in linguistics, the 'engineering-use' in computer science, and the 'experimental-use' in cognitive biology. In particular, in linguistics, the tool is developed to reveal the rules that can generate all natural languages so as to theorize the knowledge of human language (see the top diagram in Fig. 7.1). In computer science, the tool is further developed for the programmers to engineer the list of instructions such that the computer will behave according to some design needs (e.g., particular kinds of output) (see the middle diagram in Fig. 7.1). Finally, in cognitive biology, the tool is modified to probe a subject's 'internal program' through experiment (see the bottom diagram in Fig. 7.1).

On the positive side, a common target profile with an expanded usage profile in a tool offers a ‘common ground’ for the users to integrate knowledge generated from multiple disciplines. Such a common ground also allows them to bridge different, sometimes competing methodologies into one cohesive, interdisciplinary research program. For instance, Chomsky is known for not only the classification scheme that now bears his name, but also his criticism of animal model organisms used in the study of human language (Chomsky 1956a; Osgood 1975). Fitch and Hauser’s ingenious, novel application of the Chomsky hierarchy enables a ‘conceptual crossover.’ It brings nonhuman animals as model organisms back to the study of human language, together with formal models of language constructed by Chomsky.

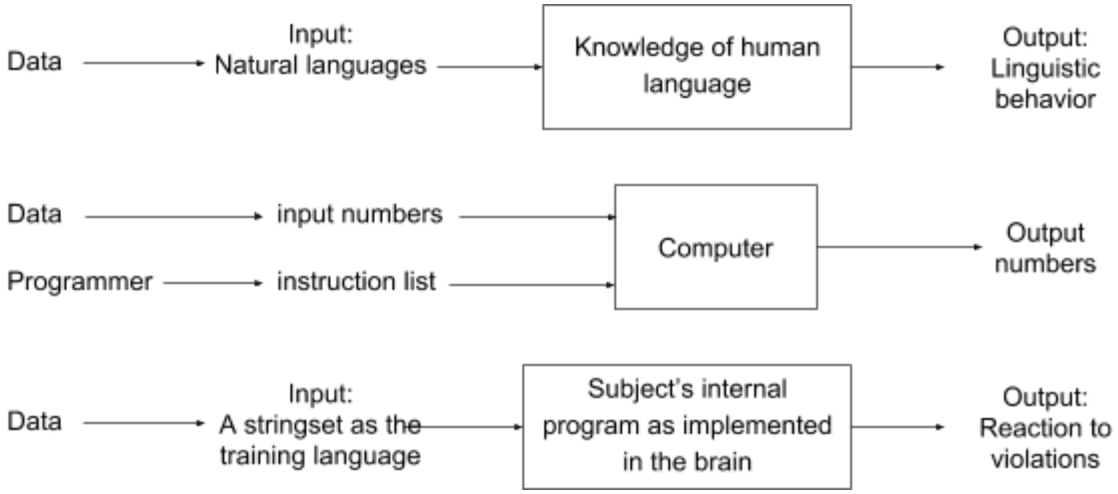


Fig. 7.1 Three Diagrams of the ‘Flow of Information’
These three diagrams illustrate different uses of formal language theory across linguistics, computer science, and cognitive Biology

On the negative side, by applying the tool in neurolinguistics, one crucial change in subject matter is left unaccounted for — the subject matter changed from being about information processing *regardless* of implementation to being about linguistic processing *actually* implemented in the brain. The lack of a counter-actor to this change of subject has caused resistance from some neurolinguists. There were also backlashes from the comparative experimental psychologists who believed that the tool could be used to detect recursion — something that goes beyond the intended use by Fitch and Hauser (2004).

A primary finding of this dissertation is that tool migration is epistemically risky, and yet the source of the epistemic risks does not necessarily come from changes to the tool. As these two illustrative case studies show, having a consistent target profile is neither necessary nor sufficient to prevent errors that are associated with tool migration.

7.2 Future Directions

7.2.1 Are Some Types of Tool Migration Safer Than Others?

In this dissertation, what I have not been able to address is whether there can be any general connections between the types of tool migration and the kinds of epistemic risks. It is possible that tool-transformation may promote counter-actors, which makes it a 'safer' type of tool migration. Conversely, it is likely that tool-application may create a 'false sense of security,' which makes it a less safe type of tool migration. For when there is a salient change in the tool's target profile, one would expect the users in the new disciplinary context to be more alert than when there isn't such a salient change. To illustrate this contrast, consider a mundane driving experience. Consider yourself driving a car that is borrowed for a long road trip. When the car is of a different brand, different year, and thus different interior from your own car, you might be more attentive to its condition before and while driving it. In contrast, when the car you borrowed happens to be of the same brand, only, say, two years older than your car, and thus it has a very similar interior to the one you have been driving for years, it would not be surprising if a sense of familiarity desensitizes you from paying attention to this car's idiosyncrasies.[53] In a sense, driving the latter car makes the road trip more dangerous than driving the former car. However, it requires further studies to identify (if there is any) connections between particular types of tool migration and specific kind of epistemic risks.

7.2.2 A New Genre of Epistemic Risks

In addition to the discussion of model transfer, another line of literature in the philosophy of science that I could make a connection with is the existing discussion of epistemic risks. The concept of epistemic risks has been varyingly associated with other ideas, such as uncertainty, probability, values, or responsibility (e.g., Sahlin 2012, Fallis 2007, Parascandola 2010, Santoro, Marino, and Tamburrini 2008, Freedman 2014). In contrast, a systematic account has been very recently proposed by Justin Biddle and Rebecca Kukla (2017, also Biddle 2016). According to them, 'epistemic risk' refers to the risk of being wrong in the context of falsely accepting or rejecting an element related to the pursuit of knowledge. Such an element could be an hypothesis, a methodology, a background assumption, a set of test subjects, a policy, or a definition. Along with this line of discussion, epistemic risks associated with tool migration can be understood as the risk of falsely accepting or rejecting a novel use of a previously established research tool. Further case studies may then be able to contribute to this literature with examples of such risks (e.g., mischaracterization, misinterpretation, misjudgment) and how to avoid

---

[53] Julia Bursten (in personal conversation) provides this example to illustrate a 'false sense of security'.

them.

7.2.3 The Unity of Science 3.0?

On top of providing preventive measures, a study of tool migration has potential to shed light on the long-standing philosophical question of how ‘unified’ science can be. Philosophers of science once aimed to reveal a logical structure by which scientific theories could be reduced to one another (e.g., Oppenheim and Putnam 1958), but this pursuit of theory reduction has long been abandoned (Kitcher 1999). Superseding this tradition was a gradual but steady shift of focus from questions concerning scientific theories to those concerning scientific practice, e.g., what counts as evidence or what counts as an explanation—a trend that persists to this day. As Humphreys (2016) argues, the focus on scientific practice has led to both a ‘disunity’ view of science among philosophers and the fragmentation of philosophy of science into various sub-disciplines, such as the philosophy of physics, the philosophy of biology, and many more. According to Humphreys, both consequences are understandable. For instance, a theory of scientific evidence that works in astronomy may not work in chemistry. Likewise, a theory of scientific explanation that accounts for explanations in biology may not necessarily account for explanations in physics or economics. Naturally, philosophers of science gravitate towards a specific science to avoid overgeneralization. Once such a division of labor is established, it becomes harder for philosophers of a specific science to grasp what may be the unifying features of science as a whole. In contrast, studying tool migration provides a different perspective for looking at scientific practice. Instead of focusing on one disciplinary context at a time. A study of tool migration compares and contrasts multiple scientific contexts, by examining how each of these disciplines handles a particular research tool. Following versatile research tools may thereby allow a common theme that underlies multiple scientific disciplines to emerge.

7.2.4 Integration of Disciplines and Tool Migration

Finally, the incident of 'conceptual crossover' in Fitch and Hauser's (2004) novel application of FLT could make an intriguing case study to contribute to the literature of the study of 'interdisciplinary.' Under the umbrella term of 'interdisciplinarity,' academic interactions across disciplinary boundaries have been studied by scholars from the humanities and social sciences (e.g., Apostel 1972; Boden 1999; Bruun, Hukkinen, Huutoniemi, and Klein 2005; Krohn 2017; Graff 2015; Lattuca 2001; Miller 1982). A major focus of this literature has been to construct typologies — especially that can be used to effectively recognize patterns or characteristics of academic activities that promote integration between disciplines or disciplinary knowledge. The philosophers of science who have engaged in this topic have focused on the interdisciplinary dynamics that prompt the creation versatile tools such as evolutionary game theory (Grüne-Yanoff 2011, 2016). In particular, Grüne-Yanoff (2011) argues that versatile tools may be seen as products of successful 'interdisciplinary exchange.' One important question, which arose from this literature of the development and application of versatile tools in science, is whether interdisciplinary exchange leads to the integration of disciplines. That is, does the process of producing a common research tool promote the cohesion of concepts and the integration of practices (e.g., explanations, ontologies, methods, and data)? Grüne-Yanoff (2016) has argued that an interdisciplinary exchange can be successful in producing a versatile tool without necessarily integrating the disciplines involved. While Grüne-Yanoff does not rule out that some interdisciplinary exchange could succeed in both, it remains a question: Which kind of dynamics may produce not only a versatile research tool but also the integration of disciplines? The tool-adaptation of FLT in cognitive biology has the potential to shed light on this question.

7.2.5 Final Remark: A Case for the Study of Tool Migration

Making novel use of a borrowed tool can be conducive to scientific progress. Yet, as I have argued, one needs to bear in mind that tool migration is not without risks. Potential errors from using borrowed tools include mischaracterization (i.e., characterizing the phenomenon in question with inappropriate assumptions), misinterpretation (i.e., using ill-fitted background contexts to interpret the result of a calculation), and misjudgment (i.e., incorrectly rejecting or accepting a novel use of a borrowed tool). Science offers abundant examples of tool migration. A proper understanding of science would be incomplete without a systematic study of tool migration. This dissertation offers a starter kit, a conceptual framework with which one may identify, classify, and study any formal tools in migration.

References

Abe, Kentaro, and Dai Watanabe. 2011. “Songbirds Possess the Spontaneous Ability to Discriminate Syntactic Rules.” *Nature Neuroscience* 14 (8): 1067–74.

Allen, C., and M. Bekoff. 1999. *Species of Mind: The Philosophy and Biology of Cognitive Ethology*. Cambridge: MIT Press.

Alexander, J McKenzie. 2009. “Evolutionary Game Theory.” *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2009/entries/game-evolutionary/.

Andrews, Kristin. 2014. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. New York: Rutledge.

Apostel, Leo. (Ed.) 1972. *Interdisciplinarity Problems of Teaching and Research in Universities.* Paris: Organization for Economic Cooperation and Development.

Axelrod, Robert. 1980a. “Effective Choice in the Prisoner’s Dilemma.” *Journal of Conflict Resolution* 24 (1): 3–25.

_____________. 1980b. “More Effective Choice in the Prisoner’s Dilemma.” *Journal of Conflict Resolution* 24 (3): 379–403.

_____________. 1984. *The Evolution of Cooperation*. Vol. 5145. New York: Basic Books.

Baars, Bernard J. 1986. *The Cognitive Revolution in Psychology*. New York: The Guilford Press.

Backus, John W, Friedrich L Bauer, Julien Green, Charles Katz, John McCarthy, P Naur, A J Perlis, H Rutishauser, K Samelson, and B Vauquois. 1960. “Report on the Algorithmic Language ALGOL 60.” *Numerische Mathematik* 2 (1): 106–36.

Bahlmann, Jörg, Ricarda I. Schubotz, and Angela D. Friederici. 2008. “Hierarchical Artificial Grammar Processing Engages Broca’s Area.” *NeuroImage* 42 (2): 525–34.

Barker-Plummer, David. 2012. “Turing Machines.” *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/sum2013/entries/turing-machine/.

Bechtel, William, George Graham, and David A Balota. 1998. *A Companion to Cognitive Science*. Malden: Blackwell.

Bekoff, Marc, Colin Allen, and Gordon M Burghardt. 2002. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge: MIT

Press.

Biddle, Justin B. 2016. “Inductive Risk, Epistemic Risk, and Overdiagnosis of Disease.” *Perspectives on Science* 24 (2): 192-205.

Biddle, Justin B., and Rebecca Kukla. 2017. “The Geography of Epistemic Risk.” in *Exploring Inductive Risk: Case Studies of Values in Science*, edited by Kevin C. Elliott and Ted Richards, 215-37. New York: Oxford University Press.

Boden, Margaret. A. 1999. “What is Interdisciplinarity.” In *Interdisciplinarity and the Organization of Knowledge in Europe*, edited by in R. Cunningham, 13-24. Luxembourg: Office for Official Publications of the European Communities.

Boysen, Sarah T., and Gary G. Berntson. 1989. “Numerical Competence in a Chimpanzee (Pan Troglodytes).” *Journal of Comparative Psychology* 103 (1): 23–31.

Brandenburger, A. M., & Nalebuff, B. J. 1995. *The Right Game: Use Game Theory to Shape Strategy*. Boston: Harvard Business Review.

Brooks, Rodney A. 1991. “Intelligence without Representation.” *Artificial Intelligence* 47 (1–3): 139–59.

Bruun, H, J. Hukkinen, K. Huutoniemi, & J. T. Klein. 2005. *Promoting Interdisciplinary Research: The Case of the Academy of Finland.* Publications of the Academy of Finland. Helsinki: Academy of Finland.

Cai, Y., Hartnett, B., Gustafsson, C., & Peccoud, J. 2007. “A syntactic model to design and verify synthetic genetic constructs derived from standard biological parts.” *Bioinformatics*, *23*(20): 2760-2767.

Carey, Susan. 2001. “Cognitive Foundations of Arithmetic: Evolution and Ontogenisis.” *Mind and Language* 16 (1): 37–55.

Casson, M. 1994. *The economics of business culture: Game theory, transaction costs, and economic performance*. New York: Oxford University Press.

Cate, Carel ten, and Kazuo Okanoya. 2012. “Revisiting the Syntactic Abilities of Non-Human Animals: Natural Vocalizations and Artificial Grammar Learning.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367 (1598): 1984–94.

Chomsky, Noam. 1956. “Three models for the description of language.” *IRE Transactions on information theory*, *2*(3): 113-124.

__________. 1959a. “A Review of B. F. Skinner’s Verbal Behavior.” *Language* 1 (1): 26–58.

__________. 1959b. “On Certain Formal Properties of Grammars.” *Information and Control*, 2: 137–167.

__________. 1963. “Formal properties of grammar.” In *Handbook of Mathematical Psychology* edited by R.D. Luce, R.R. Bush, & E. Galanter, New York: Wiley.

__________. 1975. *The Logical Structure of Linguistic Theory / Excerpted Manuscript*. New York: Plenum.

Chomsky, Noam, and George A. Miller. 1958. “Finite State Languages.” *Information and Control* 1 (2): 91–112.

Chomsky, Noam, and M.P. Marcel P Schützenberger. 1963. “The Algebraic Theory of Context-Free Languages.” *Studies in Logic and the Foundations of Mathematics* 35: 118–61.

Church, Alonzo. 1936. “A Note on the Entscheidungsproblem.” *The Journal of Symbolic Logic* 1 (01): 40–41.

Coll, A., Wilson, M. L., Gruden, K., & Peccoud, J. 2016. GenoCAD Plant Grammar to Design Plant Expression Vectors for Promoter Analysis. *Plant Synthetic Promoters: Methods and Protocols*, 1482: 219-232.

Davis, Martin. 1958. *Computability and Unsolvability*. New York: McGraw-Hill Book Company.

de Vries, Meinou H., Padraic Monaghan, Stefan Knecht, and Pienie Zwitserlood. 2008. “Syntactic Structure and Artificial Grammar Learning: The Learnability of Embedded Hierarchical Structures.” *Cognition* 107 (2): 763–74.

Dong, S., & Searls, D. B. 1994. “Gene structure prediction by linguistic methods.” *Genomics*, *23*(3): 540-551.

Evey, R. J. 1963. *The Theory and Application of Pushdown Machines*. Cambridge, MA: Rep. No. NSF-10, Harvard Comput. Laboratory.

Fallis, Don. 2007. “Attitudes Toward Epistemic Risk and the Value of Experiments.” *Studia Logica* 86 (2). Kluwer Academic Publishers: 215–46.

Fitch, W. T., & Hauser, M. D. 2004. “Computational Constraints on Syntactic Processing in a Nonhuman Primate.” *Science* 303 (5656): 377–80.

Fitch, W. Tecumseh. 2010. “Three Meanings of ‘Recursion’: Key Distinctions for Biolinguistics.” *The Evolution of Human Language*: 73–90.

________________. 2014. “Toward a Computational Framework for Cognitive Biology: Unifying Approaches from Cognitive Neuroscience and Comparative Cognition.” *Physics of Life Reviews* 11 (3): 329–64.

________________. 2018. “Bio-Linguistics: Monkeys Break Through the Syntax Barrier.” *Current Biology* 28 (12): R695–97.

Fitch, W. Tecumseh, and Angela D. Friederici. 2012. “Artificial Grammar Learning

Meets Formal Language Theory: An Overview." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (June): 1933–55.

Freedman, Karyn L. 2015. "Testimony and Epistemic Risk: The Dependence Account." *Social Epistemology* 29 (3): 251-69.

Gardner, Howard. 1985. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic books.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. 2006. "Recursive Syntactic Pattern Learning by Songbirds." *Nature*, *440*(7088): 1204-1207.

Ginsburg, Seymour. 1980. "Methods for Specifying Families of Formal Languages - Past-Present-Future." In *Formal Language Theory* edited by Ronald V. Book, 1–22. New York: Academic Press.

Ginsburg, S, and HG Rice. 1962. "Two Families of Languages Related to ALGOL." *Journal of the ACM*, 350–71.

Graff, H. 2015. *Undisciplining Knowledge: Interdisciplinarity in the Twentieth Century*. Baltimore: Johns Hopkins University Press.

Greibach, S A. 1981. "Formal Languages : Origins and Directions." *Annals of the History of Computing* 3 (1): 14-41.

Grüne-Yanoff, Till. 2011a. "Models as Products of Interdisciplinary Exchange: Evidence from Evolutionary Game Theory." *Studies in History and Philosophy of Science* Part A 42 (2): 386–97.

________________. 2011b. "Evolutionary Game Theory, Interpersonal Comparisons and Natural Selection: A Dilemma." *Biology & Philosophy* 26 (5): 637–54.

________________. 2016. "Interdisciplinary Success without Integration." *European Journal for Philosophy of Science* 6 (3): 343–60.

Hardy, M, and S Heyes. 1999. *Beginning Psychology*. Oxford: Oxford University Press.

Harris, Zellig S. 1946. "From Morpheme to Utterance." *Language* 22 (3): 161–83.

Harris, Randy Allen. 2010. "Chomsky's Other Revolution." In *Chomskyan (R) Evolutions*, edited by Douglas A. Kibbee, 237–64. John Benjamins Publishing.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* 298:1569–79.

Haywood Jr, O. G. 1954. "Military decision and game theory." *Journal of the Operations Research Society of America*, *2*(4): 365-85.

Head, Tom. 1987. "Formal language theory and DNA: an analysis of the generative capacity of specific recombinant behaviors." *Bulletin of mathematical* *biology*, *49*(6): 737-759.

Hochmann, Jean-Remy, Mahan Azadpour, and Jacques Mehler. 2008. “Do Humans Really Learn An Bn Artificial Grammars From Exemplars?” *Cognitive Science: A Multidisciplinary Journal* 32 (6): 1021–36.

Honing, Henkjan, and Willem Zuidema. 2014. “Decomposing Dendrophilia Comment on"Toward a Computational Framework for Cognitive Biology: Unifying Approaches from Cognitive Neuroscience and Comparative Cognition" by W. Tecumseh Fitch.” *Physics of Life Reviews* 11: 375–76.

Hopcroft, John E. 1987. “Computer Science: The Emergence of a Discipline.” *Communications of the ACM* 30 (March): 198–202.

Hopcroft, John E. and Jeffrey D Ullman. 1969. *Formal Languages and Their Relation to Automata*. Reading: Addison-Wesley.

Hornstein, Norbert, Howard Lasnik, Pritty Patel-Grosz, and Charles Yang (Eds.). 2018. *Syntactic Structures after 60 Years: The Impact of the Chomskyan Revolution in Linguistics*. Vol. 129. Berlin: Walter de Gruyter GmbH.

Howlett, P, and Mary S Morgan. 2010. *How Well Do Facts Travel?: The Dissemination of Reliable Knowledge*. New York: Cambridge University Press.

Humphreys, P. 2002. Computational Models. *Philosophy of Science* 69 (September): 1–27.

___________. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. New York: Oxford University Press.

___________. 2016. “Introduction: New Directions in Philosophy of Science.” In *The Oxford Handbook of Philosophy of Science*, edited by Paul Humphreys, 1:1-12. New York: Oxford University Press.

___________. 2018. “Knowledge Transfer across Scientific Disciplines.” *Studies in History and Philosophy of Science Part A*, (February): 1-8.

Hyman, Malcolm D. 2010. “Chomsky between Revolutions.” In *Chomskyan (R)Evolutions*, edited by Douglas A. Kibbee, 265–98.

Jiang, Xinjian, Tenghai Long, Weicong Cao, Junru Li, Stanislas Dehaene, and Liping Wang. 2018. “Production of Supra-Regular Spatial Sequences by Macaque Monkeys.” *Current Biology* 28 (12): 1851–1859.e4.

Kawai, Nobuyuki, and Tetsuro Matsuzawa. 2000. “Cognition: Numerical Memory Span in a Chimpanzee.” *Nature* 403 (6765): 39.

Kitcher, Philip. 1999. “Unification as a Regulative Ideal.” *Perspectives on Science* 7 (3): 337–48.

Kjeldsen, Tinne Hoff. 2001. "John von Neumann's Conception of the Minimax Theorem: A Journey through Different Mathematical Contexts." *Archive for History of Exact Sciences* 56 (1): 39–68.

Kleene, S. C. 1951. *Representation of Events in Nerve Nets and Finite Automata*. U.S. Air Force Project Rand, Research Memorandum.

__________. 1952. *Introduction to Metamathematics*. New York: Ishi Press International.

__________. 1956. "Representation of Events in Nerve Nets and Finite Automata." In *Automata Studies*, edited by Claude Shannon and John McCarthy, 3-41. Princeton: Princeton University Press.

Klein, Julie T. 2017. "Typologies of Interdisciplinarity: The Boundary Work of Definition." In *The Oxford handbook of interdisciplinarity*, edited by Robert Frodeman, 21-34. New York: Oxford University Press.

Knuuttila, Tarja, and Andrea Loettgers. 2014. "Magnets, Spins, and Neurons: The Dissemination of Model Templates across Disciplines." *The Monist* 97 (3): 280–300.

Knuuttila, Tarja, and Andrea Loettgers. 2016. "Model Templates within and between Disciplines: From Magnets to Gases–and Socio-Economic Systems." *European Journal for Philosophy of Science* 6 (3): 377–400.

Krohn, W. 2017. "Interdisciplinary Cases and Disciplinary Knowledge: Epistemic Challenges of Interdisciplinary Research." In *The Oxford handbook of interdisciplinarity*, edited by Robert Frodeman, 40-52. New York: Oxford University Press.

Kuroda, S. -Y. 1964. "Classes of Languages and Linear-Bounded Automata." *Information and Control* 7 (2): 207–23.

Landweber, Peter S. 1963. "Three Theorems on Phrase Structure Grammars of Type 1." *Information and Control* 6 (2): 131–36.

Larson, Brooke. 2017. "Long Distance Dependencies." Oxford Bibliographies Online. 2017. https://doi.org/10.1093/OBO/9780199772810-0204.

Lattuca, Lisa. 2001. *Creating Interdisciplinarity: Interdisciplinary Research and Teaching among College and University Faculty*. Nashville: Vanderbilt University Press.

Ledley, Robert Steven. 1962. *Programming and Utilizing Digital Computers*. New York: McGraw-Hill.

Levelt, Willem J. M. 2008. *An Introduction to the Theory of Formal Languages and Automata*. Amsterdam: John Benjamins Publishing Company.

Lewis, Shevaun, and Colin Phillips. 2015. "Aligning Grammatical Theories and

Language Processing Models." *Journal of Psycholinguistic Research* 44 (1): 27–46.

Lidz, Jeffrey. 2018. "The Explanatory Power of Linguistic Theory." In *Syntactic Structures after 60 Years: The Impact of the Chomskyan Revolution in Linguistics*, edited by Norbert Hornstein, Howard Lasnik, Pritty Patel-Grosz, and Charles Yang. Berlin: Walter de Gruyter GmbH & Co KG.

Marcus, Gary F. 2006. "Startling Starlings." *Nature* 440 (7088): 1117–18.

Matsuzawa, Tetsuro. 1985. "Use of Numbers by a Chimpanzee." *Nature* 315 (6014): 57.

Maxwell, J. C. 1861. Xxv. on Physical Lines of Force: Part I.–the Theory of Molecular Vortices Applied to Magnetic Phenomena. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21 (139): 161–75.

Maynard-Smith, J. 1997. "John Maynard Smith - Scientist - The Point of Evolutionary Game Theory - Web of Stories." Video. https://www.webofstories.com/play/john.maynard.smith/50.

Maynard-Smith, J, & Price, G. R. 1972. "The Logic of Animal Conflict." *Nature* 246 (5427): 15–18.

McCulloch, Warren S, and Walter Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *The Bulletin of Mathematical Biophysics* 5 (4): 115–33.

Miller, George A. 1956. "The Magical Number Seven, plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63 (2): 81.

Miller, George A. 1967. "Project Grammarama." In *Psychology of Communication*, 125–87. New York: Penguin Books.

Miller, Raymond. 1982. Varieties of Interdisciplinary Approaches in the Social Sciences. *Issues in Integrative Studies*, 1: l-37.

Moll, Robert N., Michael A. Arbib, and A. J. Kfoury. 1988. *An Introduction to Formal Language Theory*. New York: Springer.

Morgan, Mary. S. 2010. "Traveling Facts." In *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, edited by Peter Howlett and Mary S. Morgan, 3-41. New York: Cambridge University Press.

_____________. 2014. "Resituating Knowledge: Generic Strategies and Case Studies." *Philosophy of Science*, *81*(5), 1012-1024.

Newell, Allen, John C Shaw, and Herbert A Simon. 1959. "Report on a General Problem Solving Program." In *IFIP Congress*, 256:64.

O'Regan, Gerard. 2013. *Giants of Computing*. London: Springer.

Oettinger, Anthony G. 1961. "Automatic Syntactic Analysis and the Pushdown Store." In *Structure of Language and its Mathematical Aspects*, Proceedings of Symposia in

Applied Mathematics, vol. XII, 104-29. Providence: American Mathematical Society.

Oppenheim, P. and H. Putnam. 1958. “The unity of science as a working hypothesis”, in H. Feigl et al. (eds.), *Minnesota Studies in the Philosophy of Science*, vol. 2, Minneapolis: Minnesota University Press.

Osgood, Charles E. 1975. “A Dinosaur Caper: Psycholinguistics Past, Present, and Future.” *Annals of the New York Academy of Sciences* 263: 16–26.

Parascandola, Mark. 2010. “Epistemic Risk: Empirical Science and the Fear of Being Wrong.” *Law, Probability and Risk* 9 (3–4): 201–14.

Parkes, Alan. 2002. *Introduction to Languages, Machines and Logic : Computable Languages, Abstract Machines and Formal Logic*. London: Springer-Verlag.

Partee, Barbara H., Alice G. ter Meulen, and Robert Wall. 1990. *Mathematical methods in linguistics*. Dordrecht: Kluwer Academic Publishers.

Paun, G, G Rozenberg, and A Salomaa. 2005. *DNA Computing: New Computing Paradigms*. Berlin: Springer.

Perruchet, Pierre, and Arnaud Rey. 2005. “Does the Mastery of Center-Embedded Linguistic Structures Distinguish Humans from Nonhuman Primates?” *Psychonomic Bulletin and Review* 12 (2): 307–13.

Poibeau, Thierry. 2017. *Machine Translation*. Cambridge: MIT Press.

Post, Emil L. 1943. “Formal Reductions of the General Combinatorial Decision Problem.” *American Journal of Mathematics* 65 (2): 197–215.

__________. 1944. “Recursively Enumerable Sets of Positive Integers and Their Decision Problems.” *Bulletin of the American Mathematical Society* 50 (5): 218–316.

__________. 1947. “Recursive Unsolvability of a Problem of Thue.” *The Journal of Symbolic Logic* 12 (1): 1–11.

Poundstone, William. 1993. *Prisoner’s Dilemma/John von Neumann, Game Theory and the Puzzle of the Bomb*. New York: Random House.

Premack, David. 2004. “Is Language the Key to Human Intelligence?” *Science* 303 (5656): 318–20.

Pullum, Geoffrey K. 2011. “On the Mathematical Foundations of ‘Syntactic Structures.’” *Journal of Logic, Language, and Information* 20 (3): 277–96.

Putnam, Hilary. 1961. “Some Issues in the Theory of Grammar.” *Structure of Language and Its Mathematical Aspects* 12 (20): 25–42.

Rey, Arnaud, Pierre Perruchet, and Joël Fagot. 2012. “Centre-Embedded Structures Are a

By-Product of Associative Learning and Working Memory Constraints: Evidence from Baboons (Papio Papio).” *Cognition* 123 (1): 180–84.

Sahlin, Nils-Eric. 2012. “Unreliable Probabilities, Paradoxes, and Epistemic Risks.” In *Handbook of Risk Theory*, 477–98. Dordrecht: Springer Netherlands.

Sammet, Jean E. 1972. “Programming Languages: History and Future.” *Communications of the ACM* 15 (7): 601–10.

Santoro, Matteo, Dante Marino, and Guglielmo Tamburrini. 2008. “Learning Robots Interacting with Humans: From Epistemic Risk to Responsibility.” *AI & SOCIETY* 22 (3): 301–14.

Savage, John E. 1998. *Models of Computation: Exploring the Power of Computing*. Reading: Addison-Wesley.

Schneider, Lambert. 2010. “A Journey through Times and Cultures? Ancient Greek Forms in American Nineteenth-Century Architecture.” In *How Well Do Facts Travel?: The Dissemination of Reliable Knowledge*, edited by Peter Howlett and Mary S. Morgan. New York: Cambridge University Press.

Schubert, Lenhart. 2019. “Computational Linguistics.” *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/spr2019/entries/computational-linguistics/.

Searls, David B. 1992. “The linguistics of DNA.” *American Scientist*, *80*(6): 579-591.

__________. 1993. “The computational linguistics of biological sequences.” *Artificial intelligence and molecular biology*, *2*: 47-120.

__________. 1995. “String variable grammar: A logic grammar formalism for the biological language of DNA.” *The Journal of Logic Programming*, *24*(1-2): 73-102.

__________. 2002. “The Language of Genes.” *Nature* 420 (November): 211.

Schützenberger, Marcel Paul. 1961. “On the Definition of a Family of Automata.” *Information and Control* 4 (2–3): 245–70.

Skyrms, Brian. 2010. *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.

Snidal, Duncan. 1985. “The Game Theory of International Politics.” *World Politics*, *38*(1): 25-57.

Stobbe, N., G. Westphal-Fitch, U. Aust, and W. T. Fitch. 2012. “Visual Artificial Grammar Learning: Comparative Research on Humans, Kea (Nestor Notabilis) and Pigeons (Columba Livia).” *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1598): 1995–2006.

Taylor, Peter D., and Leo B. Jonker. 1978. “Evolutionary Stable Strategies and Game Dynamics.” *Mathematical Biosciences* 40 (1–2): 145–56.

Turing, A. M. 1936. "On Computable Numbers, with an Application to the Entscheidungsproblem." In *Proceedings of the London Mathematical Society*, 2:230–65.

Udden, Julia, Martin Ingvar, Peter Hagoort, and Karl M Petersson. 2012. "Implicit Acquisition of Grammars with Crossed and Nested Non−adjacent Dependencies: Investigating the Push−down Stack Model." *Cognitive Science* 36 (6): 1078–1101.

van Heijningen, Caroline A A, Jos de Visser, Willem Zuidema, and Carel ten Cate. 2009. "Simple Rules Can Explain Discrimination of Putative Recursive Syntactic Structures by a Songbird Species." *Proceedings of the National Academy of Sciences of the United States of America* 106 (48): 20538–43.

von Neumann, John. 1928. "Zur Theorie Der Gesellschaftsspiele (On the Theory of Games of Strategy)." *Mathematische Annalen*, 295–320.

________________. 1959 "On the theory of games of strategy." *Contributions to the Theory of Games* 4, 13-42.

von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. London: Oxford University Press.

Wells, Rulon S. 1947. "Immediate Constituents." *Language* 23 (2): 81–117.