Literacy, linguistics and compositionality: Investigating the effects of cultural systems on learning and language.

Erin Brown, B.I.S.

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

August 2008

Submitted in partial fulfilment of the requirements of the degree of Master of Science.
Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Erin Brown
Acknowledgements

I would like to thank the following people for support during this dissertation:

First, to my supervisor, Monica Tamariz, whose expert guidance and creative insights made this project possible. Also to my co-supervisor, Simon Kirby, for his invaluable advice and much-needed musical knowledge.

To Eddie Dubourg, for his patience and help with Psycope and audio software.

To Chrissy and James, whose camaraderie and masterful sushi skills made this experience fun.

I am also extremely grateful to my parents, who always support me in any way they can. I thank my sister, Kara, for sharing her hard-earned wisdom and unsurpassed knowledge of English grammar. I also thank Courtney for coming all the way across the Atlantic and being such a great distraction. Finally, I thank Josh for his interest, enthusiasm and encouragement.
Abstract

Recent linguistic research has shown that cultural processes operating over an extended timescale may be responsible for many aspects of syntax. Other evidence from artificial language learning studies indicates a strong bias for systematicity, potentially conflicting with these cultural accounts of language structure. However, such studies do not consider the influence of literacy, which may have been a primary determinant of their participants' behaviour. This dissertation explores the cognitive and academic consequences of literacy and empirically investigates the influence of cultural systems on language learning in a different modality, that of music. Musical literacy is shown to induce a similar bias for systematicity to orthographic literacy in traditional tasks. The trends found indicate that access to abstract representations alters learning. This work demonstrates the need to consider literacy when undertaking linguistic research, particularly the study of language evolution, as an accurate and complete understanding of language is a prerequisite for explaining its origin.
## Contents

Chapter 1  Introduction ................................................................. 7

Chapter 2  Background ................................................................. 9  
  2.1 Culture, cognition and language ........................................ 9  
     2.1.1 Summary ............................................................... 10  
  2.2 Compositionality and culture ........................................... 10  
     2.2.1 Computational modelling ....................................... 11  
     2.2.2 Socio-cultural factors .......................................... 12  
     2.2.3 Artificial language learning ................................... 13  
     2.2.3.1 Discussion ..................................................... 16

Chapter 3  Literacy and language ................................................. 18  
  3.1 Literacy and the individual ............................................. 18  
     3.1.1 Phonology ......................................................... 18  
     3.1.2 Morphology, semantics and the lexicon ...................... 19  
     3.1.3 Syntax ............................................................... 20  
  3.2 Literacy and languages .................................................. 21  
  3.3 Discussion ................................................................. 22  
     3.3.1 Summary ............................................................. 22  
     3.3.2 Consequences for linguistics .................................. 23

Chapter 4  Language Evolution ...................................................... 25  
  4.1 Characterising language .................................................. 25  
     4.1.1 Literacy and linguistic uniformitarianism .................. 25  
     4.1.2 Syntax as a by-product of symbolic relationships ...... 26  
     4.1.3 Child language acquisition .................................... 27  
     4.1.4 Ape language research ......................................... 28  
     4.1.5 Summary ............................................................. 30  
  4.2 Language origins ........................................................... 30  
  4.3 Discussion ................................................................. 32

Chapter 5  An Alternative Approach ............................................. 34  
  5.1 Musical language learning ............................................... 35  
     5.1.1 Components of a musical language ......................... 36  
     5.1.2 Measuring learning .............................................. 37

Chapter 6  The Present Experiment .............................................. 40  
  6.1 Methods ..................................................................... 40  
     6.1.1 Counteracting unrelated advantages from musical training ... 41  
     6.1.2 Method of analysis .............................................. 42  
     6.1.3 Materials ............................................................. 43  
     6.1.4 Procedure ............................................................ 44  
  6.2 Results and discussion ................................................... 46  
     6.2.1 Compositionality of individual languages ................ 47  
     6.2.1.1 Discussion ..................................................... 47  
     6.2.2 Structure of output compared to input ..................... 50  
     6.2.2.1 Discussion ..................................................... 51  
     6.2.3 Similarity of form .............................................. 54  
     6.2.3.1 Discussion ..................................................... 55

Chapter 7  General Discussion .................................................... 58  
  7.1 Summary of results ....................................................... 58
Contents (cont.)

7.2 Meta-musical awareness and evidence for the systematicity bias ..................... 59
7.3 Broader implications .......................................................................................... 60

Chapter 8 Conclusion ......................................................................................... 62
Appendix I Input Languages .................................................................................. 63
Appendix II Questionnaire .................................................................................... 64
References ........................................................................................................... 65
Chapter 1 Introduction

The history of linguistics is characterised by radical shifts in thinking. The extreme cultural relativism of early anthropological accounts (Whorf, 1956) and the inadequacy of behaviourist theories (Skinner, 1957) created an intellectual backlash (Chomsky, 1959) that emphasised the universal properties of language. This perspective remains prevalent in theoretical linguistics, but a very different characterisation of language has emerged from the development of new disciplines. While generative grammar proponents focus on defining a formal set of rules for describing language structure, pragmatists, sociolinguists and others study the dynamics of language in use. This more functional approach views language not as a rigid rule system, but an elusive and malleable entity, greatly affected by pressures arising from human communication. Concomitant with these differing perspectives on structure are equally opposing ideas concerning its origin. Those viewing language as a formal grammar generally point to the operation of innate properties of the human brain, while others consider culture to be the dominant force.

The highly divergent opinions within linguistics today can in part be attributed to the different sources of data upon which theories are based. Written language typically serves as primary data for traditional grammar theories, while speech is often the focus of social approaches to language. That major differences exist between the two is widely acknowledged, but the full extent of these is rarely appreciated, despite the fact that the implications have considerable consequences for most areas of linguistics. The available research suggests that relying on written sources has skewed theorists’ understanding of language towards a strict, rule-governed system. Syntax is central to these theories, but its significance may have been inflated by erroneous assumptions. The fact that systematic relationships are less rigid in speech than in writing brings into question the need for an innate endowment of linguistic knowledge, thus strengthening explanations based on cultural processes. However, precisely how and to what extent culture influences linguistic phenomena remains to be sufficiently understood.

The functions of learning and experience in the development of syntactic competence must be explored in a more interdisciplinary context, considering research on the
nature of language itself, the acquisition process and the dynamics of cultural processes. The effects of cultural systems, in particular literacy, on both individual linguistic behaviour and the field of linguistics must be thoroughly understood before accurate explanations of language structure can be formulated.

This paper will first review research concerning the role of culture in shaping language structure, focusing particularly on compositionality, a concept encompassing many aspects of syntactic relationships. The possibility that cultural processes are primarily responsible for grammatical structure will be considered in Chapter Two, and artificial language learning studies undertaken to address this issue will be discussed in detail. Chapter Three will address the impact of literacy on language and cognition, as well as the development of theories of grammar. In Chapter Four perspectives on language evolution will be discussed in light of the evidence presented. The remainder of the dissertation will present an experiment designed to elucidate issues regarding systematicity and cultural knowledge as they interact in artificial language learning tasks. The findings will be discussed within the broader context of evolutionary linguistics.
Chapter 2  Background

2.1 Culture, cognition and language
The issue of nature versus nurture in human cognitive development is so difficult to resolve because the interaction of innate mechanisms and experience is immensely complex, beginning while a child is still in the womb (DeCasper & Spence, 1987). Recent research has demonstrated that cognition can be shaped in subtle but observable ways by prior experience (Kay et al., 2007; Boroditsky, 2003). The fact that behaviour can be significantly altered by learning highlights the critical function it serves in cognitive development. This also suggests that the acquisition of language may be heavily influenced by experience, and therefore makes determining the role of environment and culture in this process all the more important.

A prominent position in linguistics holds that many aspects of language arise from innate, biological mechanisms (Pinker, 1994). Chomsky (1965) first proposed that children must innately possess abstract linguistic knowledge, usually referred to as Universal Grammar (UG), that aids them in acquiring language, an idea that has received wide support throughout linguistics (Bickerton, 1990; Pinker, 1994; Jackendoff, 2002). Language structure is regarded as so exceedingly complex that a biological endowment must exist to account for humans’ remarkable ability to learn and use language effortlessly. Implicit in this view of acquisition is the idea that language consists of a formal set of rules for systematically constructing words, phrases and sentences.

Despite the prominence of innatist perspectives, alternative theories of both language and the acquisition process have challenged the traditional view (Langacker, 1987; Tomasello, 2000; MacWhinney, 2004). Recent empirical research in developmental psychology and cognitive science brings into question the assumptions of traditional approaches and the concept of natural language as a formal language (Tomasello, 2003a). Tomasello (2003a; see also Schoenemann, 1999; Deacon, 1997) stresses the importance of meaning and symbolic relationships in determining linguistic structure. Linguistic knowledge is represented not by rules but a ‘structured inventory of symbolic units’ (pg. 105). Language structure, it follows, arises from language use.
In this view, cultural processes create language structure, not information directly encoded in the genome.

Culture is also strongly implicated in language acquisition. Tomasello (2003a) contends that function, meaning and context provide sufficient aid for children in the acquisition process. He points to strong correlations between the content of children’s language and the speech of their caregivers, which suggests the linguistic environment plays a major role in grammar acquisition. In addition, he argues that general socio-cognitive skills, such as pattern analysis and cultural learning, are employed for learning language, negating the need for innate, specifically linguistic, knowledge.

2.1.1 Summary
The research reviewed offers strong support for shifting the emphasis in an explanation of language from the biological to the cultural. Although Tomasello (2003a) considers grammar a product of cultural processes, he only provides a few modern examples of grammaticalisation as illustration. He does not discuss how a language could have begun in a qualitatively different form and gradually become, through cultural processes alone, the highly structured systems observed today. A number of authors have addressed this issue directly in recent years (Cornish, 2006; Hurford & Kirby, 2002; Kirby, 2000; Hurford, 1998; Batali, 1998). The following section will focus on research that suggests a central component of syntax may not be a product of human biology, but instead a result of cultural evolution.

2.2 Compositionality and culture
Compositionality, a defining characteristic of languages spoken today, allows speakers to create novel utterances to convey novel meanings that other language users will understand (Hockett’s (1969) ‘productivity’). It is a particular type of systematicity, in which the meaning of a whole is a function of both its parts and the way in which they are combined. Meaning is transparent in compositional structures, as a whole meaning can be discovered through analysis of its parts. Perhaps because of its exceptional effectiveness in conveying meaning and critical role in languages today, compositionality is often thought to be part of a biological component of human cognition, one that both directly causes language structure and is required for
children to successfully acquire grammars (Pinker & Bloom, 1990; Pinker, 1994). Recent work challenges this view, however, and suggests that the biological aspects of grammar may be much less substantial than previously thought.

It has recently been argued that the cultural evolution of language is a Darwinian process (Mesoudi et al., 2006), and emergent phenomena are very likely to result from the cultural processes operating during its acquisition and use. On a historical timescale, the accumulation of very small changes ultimately leads to major shifts. From the perspective of an individual, the effects of these changes will not be apparent; however, from the perspective of a population of individuals interacting over time, larger patterns can be observed. It is precisely this process that computational simulations attempt to capture, by isolating one or a few component mechanisms to understand better their distinct contributions.

### 2.2.1 Computational modelling

Computational modelling is a method for studying language that allows researchers to investigate highly complex and dynamic processes, the outcomes of which very often contradict our intuitions (Bechtel & Abrahamsen, 1991). The aim of many simulations has been to discover what types of patterns emerge based on cultural evolution alone (Brighton, 2002; Kirby, 2002; Smith, 2002; Brighton & Kirby, 2001; Batali, 1998).

Batali (1998) investigated what type of communication system arises within a population of agents with no initial grammar. Simply through repeated exchange and statistical fluctuations in signal production and interpretation, the population eventually converged on a (predominantly) shared communication system. Not only did a coordinated system develop, but the structure of its signals appears to be partially compositional. Although it is not immediately obvious why this outcome occurs, it is nonetheless striking that such a highly structured system emerges through cultural processes alone.

Kirby & Hurford (2002) discuss how language evolution involves the interaction of three separate, adaptive systems (individual development, biological evolution and cultural evolution) operating on separate but overlapping timescales, therefore making
it extremely difficult to tease apart the independent effects of each. The Iterated Learning Model (ILM) was developed to address these issues (Kirby, 2000). The distinguishing feature of the ILM is cultural transmission over generations, with the output of one serving as the input for the next. The initial agents in simulations possess no grammar and produce random strings of characters to convey meanings. With only simple heuristic strategies available, over thousands of generations of repeated learning and production, fully compositional systems emerge. This only occurs, however, when agents observe a limited portion of the meaning-signal pairs in a language, what Kirby terms a ‘learning bottleneck’. He likens this to the so-called ‘poverty of the stimulus’ - the argument that input from the linguistic environment is insufficient for explaining grammar acquisition – that is widely invoked in arguments for the innateness of grammar (Pullum & Scholz, 2002). Based on these results, Kirby claims that compositionality may in fact be a consequence of purely cultural processes operating over an extended timescale.

The fact that defining characteristics of language can emerge in systems through cultural processes alone suggests it may not be necessary or correct to infer a direct causal link between biology and language structure (Kirby & Hurford, 2002). Computational research sheds light on how dynamic systems operate and demonstrates that structure can emerge under certain conditions; however, without empirical verification we cannot extend the findings to human language. Language change in human communities must also be examined in order to better understand how the social forces of language use ultimately affect language form.

2.2.2 Socio-cultural factors
The development of sociolinguistics (Labov, 1966) and pragmatics (Grice, 1975; Brown & Levinson, 1978; Sperber & Wilson, 1986) marked a greater appreciation of the social and cultural aspects of language. Knowledge that social factors can strongly influence linguistic behaviour brought into question the rigidity of traditional characterisations of language. Although concepts as central as grammaticality were shown to be neither simple nor straightforward, this drew little acknowledgement from linguists working on theories of grammar (Schoenemann, 1999; Linell, 1982). Even today, many linguists choose not to consider the ‘messy’ side of language embodied in speech and casual communication, and instead continue to focus on an
idealised form. However, some authors have commented on the far-reaching consequences the implications of this research have for many areas of linguistics (Linell, 1982; Newmeyer, 2002; Tomasello, 2003b).

Western linguists tend to apply grammatical categories of European languages onto other languages of the world, leading to analyses that may not accurately or usefully reflect the structure of these languages (Everett, 2005; Grace, 2002). This practise is likely in part an attempt to defend the concept of language universals (Tomasello, 2003a). That all grammars share some features is not disputed, but specific grammatical constructions do not seem to be among these (Tomasello, 2003a). True universals turn out to be very abstract qualities, which are probably mere consequences of the complex symbolic relationships represented in our internal concepts (Schoenemann, 1999).

Acknowledging that languages around the world may differ radically in structure is an important step, but exactly how and why particular forms or grammatical strategies are present in some languages and not others requires explanation. Wray & Grace (2005) observe a distributional pattern of structures in the world’s languages and argue that socio-cultural factors can account for much of this variation. The authors make a distinction between exoteric and esoteric languages. Esoteric languages are used for communication between people with shared cultural identities and environments, relying heavily on context and shared knowledge. Exoteric communication, on the other hand, occurs between strangers or members of different social groups and requires that most of the informational content of a message be stated explicitly. Esoteric languages, typically only learned by children as a native language, exhibit features that create a barrier to communication with outsiders, such as complex phonological systems with unusual sounds, complex morpho-phonological rules with difficult to pronounce combinations, highly irregular morphology and a large inventory of holistic (non-systematic) constructions. Exoteric languages are more frequently learned by adults as a second language and used for a wider variety of purposes, often in the textual modality, so novel meanings must be conveyed in a way that will be easily understood. This type of communication requires systematic relationships between forms and meanings, and thus leads languages to utilise more compositional constructions.
Wray & Grace (2005) claim that the socio-cultural dimensions of language use can influence language structure, and recent computational modelling research supports their claims (Lupyan & Dale, 2008). Wray & Grace also review various linguistic strategies available to speakers and the different circumstances under which these are employed, concluding the default state of language is more likely that of esoteric languages. They suggest that languages of the distant past may have been markedly different from those spoken today, speculating that humans did not always utilise the highly systematic strategy prevalent in some languages presently. Analytic, or systematic, structuring is a possibility for satisfying communicative needs, which has arisen, and in some cases become prominent, due to socio-cultural processes. This corroborates Kirby’s (2000; 2002) explanation of compositionality as resulting from cultural forces.

Repeated learning and cultural evolution influenced by the social dynamics of language use appear to be major factors shaping language structure. However, the possibility that an innate property of human cognition is the source of compositionality in natural languages has not yet been ruled out. Wray and Grace’s (2005) proposal is based on observed patterns in a survey of the world’s languages in conjunction with psychological research on processing pressures and language acquisition. That particular linguistic phenomena correlate strongly with cultural factors does not definitively prove that these factors are the cause of highly structured systems. It is still necessary to approach the issue within an empirical framework. Directly observing the linguistic behaviour of speakers allows us to understand how individual language learning and use ultimately affect systems of communication.

2.2.3 Artificial language learning

Artificial language learning (ALL) studies have examined many aspects of linguistic behaviour as they occur spontaneously in human learners (Kaschak & Saffran, 2006; Hudson-Kam & Newport, 2005; Christiansen and Ellefson, 2002; Gómez & Gerken, 2000; Esper, 1966). The paradigm involves training participants on an artificially-constructed, miniature language and examining what and how much they have learned. ALL studies can reveal, among other things, learning strategies, how
linguistic systems are internally represented, constraints on learning and how much of what has been learned is implicitly or explicitly known.

A recent study adapted the ILM for use with human learners using an ALL paradigm (Cornish, 2006). Participants were trained on languages consisting of paired objects and words, with objects presented as images and words presented in text. Parameters of variation were the exact proportion of a learning bottleneck (50% or 75%) and structure of the initial language (fully compositional or random). After training, participants were required to provide words for the full set of meanings. With a fully compositional initial language and a 75% bottleneck, overall structure of the language remained significantly high throughout the generations. For the most part, participants successfully learned and reproduced their input languages. When the initial language was random, final languages for both bottleneck conditions were compositional to only a small degree. Even so, these still exhibited some structure, and inspection of intermediate phases revealed that at certain points in the chain of participants overall structure had become quite high.

This study demonstrates that iterated learning and the opportunity to generalise can lead learners to impose systematicity on a language, and also that highly systematic languages are learned and reproduced to a greater degree than random ones. One drawback of this approach is the problem of individual variation. Some participants are sensitive to structure and will either accurately learn existing systematic relationships or innovate their own. Others seem to be less aware of structural properties, producing output that does not contribute to the gradual increase in systematicity, at times greatly reducing it.

Another ALL study addressed similar issues but attempted to control for the effects of individual variation by excluding iterated learning (Tamariz & Smith, submitted). This design also more closely examines individual learning and reproduction processes, where innovation must ultimately occur. A new method for measuring systematicity, RegMap, was used. RegMap provides a more thorough and descriptive

---

1 RegMap as a measurement method differs in a subtle but important way from that of Cornish (2006). Cornish measured structure by finding the correlation between similarities in the meaning space and similarities in the signal space, and measured how well a language was learned by finding the distances between signals in the input and output languages. RegMap, instead, is based on co-occurrence
quantification of mappings between meanings and signals, both for individual dimensions and a language as a whole.\(^2\)

In this design, four groups of participants were trained and tested on the full meaning-signal set of four different languages with varying levels of compositionality (from mostly random to fully compositional). Learnability was measured as the correlation between the systematicities of participants’ input and output languages, and an interesting pattern between groups was discovered. For participants trained on partially compositional input, some produced more compositional outputs. In addition, most participants trained on a fully compositional language produced fully compositional output.

These results, the authors claim, are evidence for what they term a systematicity bias, or a ‘cognitive bias in favour of learning and reproducing compositional patterns in symbolic systems’ (pg. 16). The fully compositional system was reproduced accurately in the majority of cases, and some participants in other conditions imposed regularities on somewhat irregular systems. Though the results indicate the presence of a systematicity bias, its origin remains unknown. The authors speculate that it could be innate, learned or a consequence of socio-cultural factors. They cite literacy as a potential cause, implying that the bias is not innate.

2.2.3.1 Discussion
Both of the ALL studies discussed provide evidence for the systematicity bias reported by Tamariz & Smith (submitted). As observed, it would appear to be a particularly strong bias. Systematisation through iterated learning took place over merely ten generations of individual learners (Cornish, 2006). In Tamariz & Smith (submitted), the effects of the bias were evident after only one instance of learning. Although Cornish proposes that her results support cultural explanations of language structure, it is important to note the discrepancy in timescales. According to culture-

\(^2\) For the remainder of this paper, the term systematicity will be used for the quantified structure of an individual meaning-signal dimension and compositionality for structure of whole languages.
based theories, changes in linguistic structure are minute and rare, the overall effects of which only become apparent after hundreds or thousands of years. The observed systematicity bias, however, would cause a similar outcome in a much shorter length of time. Cultural processes appear to be less important than the processes in operation during individual learning and production. The nature and strength of the bias make it more amenable to interpretation as an innate feature of human cognition, contradicting the idea that structure arises from cultural transmission.

These conflicting accounts may be resolved by pinpointing the source of the systematicity bias, with the aim of discovering how much can be attributed to prior learning. Surely knowledge of the highly structured languages already possessed by participants could itself cause them to seek out or apply regular patterns. It is also possible that their experience with written language greatly enhances this knowledge, as hypothesised by Tamariz & Smith (submitted). Written language often makes meaning-signal relationships regular and explicit. For instance, the English regular past tense morpheme has three different forms in speech (/-d/, /-t/ or /-ed/), but is usually represented orthographically in all three instances as –ed. This explicit regularity likely increases an awareness of and expectation for systematicity in language.

In order to determine the role of culture and learning in shaping linguistic behaviour, and thus its effect on language structure, this hypothesis requires further investigation. If the rapid regularisation observed in these studies can be shown to result from knowledge of a cultural system, it would provide support for the gradual cultural accounts of language evolution. Although systematicity in its various forms seems to be a defining and central feature of language, its status must first be considered more carefully, specifically its relation to literacy. The nature of compositionality in language can be better understood by examining how written language has affected the development of linguistic theories, individual cognitive and linguistic development, and languages on a larger social scale.
Chapter 3  *Literacy, Cognition and Language*

Any linguist will acknowledge that orthography is not a core feature of language. Written language is not present in all cultures and is of relatively recent origin, the earliest forms appearing approximately 5,000 years ago (Schmandt-Besserat, 1996). Literacy can have surprising and interesting effects on cognition, such as altering the conceptualisation of abstract categories (Ong, 1982) and significantly enhancing the functional lateralisation of language to the left hemisphere (Petersson et al., 2007). Even so, the impact of literacy on both cognition at an individual level as well as languages and societies is often ignored or not well understood by those developing linguistic theories (Linell, 1982; Grace, 2003). Although much psychological and sociological research has been devoted to understanding literacy, it has not been sufficiently incorporated into the field of linguistics.

If literacy can significantly influence brain organisation and more general aspects of cognition, its role in shaping linguistic behaviour must be particularly profound. On the level of the individual, its impact spans multiple modalities of language. In addition, widespread literacy and the easy availability of texts can play a central role in language change as it takes places in societies. This Chapter will discuss how literacy affects the individual and languages as entities themselves.

3.1 Literacy and the individual

Psychological research on literacy and its relationship to metalinguistic knowledge has been subject to conflicting interpretations, most significantly in establishing a causal relationship between the two (Olson, 1996). The following sections will discuss the effects of literacy on different aspects of linguistic behaviour and knowledge, and also how these findings should be interpreted.

3.1.1 Phonology

Learning to read and write is associated with a marked increase in phonological awareness (Ehri, 1985). Early researchers proposed that children’s pre-existing phonological awareness aids them in becoming literate, and thus viewed literacy skills as based upon prior ability (Nunes et al., 2006). However, later authors questioned
the assumption that a sufficient level of metalinguistic knowledge is present before training in orthography and posited that literacy itself induces phonological awareness (Olson, 1996). In other words, the emergence of explicit knowledge of the structural units of speech will not occur without experience with a written system for representing those sounds. These claims are supported by research with illiterates, whose behaviour mirrors preliterate children on similar tests (Morais et al., 1979).

Importantly, orthography influences speech processing even in the absence of text. This is evident from spelling pronunciations in aural tasks (Ehri, 1985) and the errors observed in aural rhyming tasks (Ziegler & Ferrand, 1998; Taft & Hambly, 1985). In addition, illiterates’ performance is much lower than literates’ in pseudoword repetition tasks, and this difference is evident in a functional analysis of brain activity and organisation (Castro-Caldas et al., 1998; Petersson et al., 2000).

The evidence reviewed suggests that conscious knowledge of the structural units of speech and the ability to manipulate them is a consequence of acquiring a written system of representation. Olson (1996) compares the process of becoming literate to the gradual, historical development of alphabetic scripts, which he characterises as a continuing discovery of linguistic structures. He claims that early graphic symbols enabled humans to create abstract representations of spoken language, and through historical and cultural processes, these were slowly broken down into individual sound units. Olson provides further support for phonological awareness arising from experience with orthography, noting that people are unaware of those aspects of speech not represented in writing, or claim to perceive units that appear in writing but not speech. To be sure, the phonological structure of language is accessible without the aid of orthography (otherwise, scripts could not have been invented in the first place). However, in the absence of orthography, phonological awareness is much less likely to develop (Ehri, 1985).

3.1.2 Morphology, semantics and the lexicon
Not only are units of sound made explicit in writing, but so are units of meaning, particularly in an orthographic system like written English that often marks morphological units with spellings that contradict regular letter-to-sound conventions (e.g. sign and signal). Spelling ability in children predicts morphological awareness,
indicating that learning to read and write increases awareness of the meaningful units of language (Nunes et al., 2006). Thus, just as the phonological structure of language becomes apparent to the literate user, so does the morphological.

Verbal working memory is also enhanced by literacy (Peterssson et al., 2000), and orthography can increase memory in word learning and recall (Ehri & Wilce, 1979; Sales et al., 1969, cited in Ehri, 1985). This point is critical to the present discussion concerning compositionality in ALL tasks and will be returned to in section 3.3.1. Literates’ and illiterates’ performance on semantic verbal fluency tests differ both qualitatively and quantitatively depending upon the task. When asked to list items from a category that is part of everyday life, both groups perform similarly; however, if the category is more abstract and not part of personal experience, illiterates list significantly fewer items than literates (Gonzalez da Silva et al., 2004).

Linguists and most literate language users would probably agree that the word is the fundamental unit of speech. However, the concept of words as discrete and isolated units is not necessarily shared by preliterate speakers (Ong, 1982), and it has been argued that the fundamental unit of speech is in fact the ‘idea unit’ (Chafe, 1985). Although children must have some understanding of the distinction between content words in speech in order to learn what individual words refer to, function words are treated in a very different manner before literacy is acquired. Preliterate children have difficulty isolating single words from adjacent ones in sentences, especially function words like should and for, which carry little meaning out of context (Ehri, 1985). In some cases, when presented alone, they are not even identified as real words (Ehri, 1985). There is evidence from multiple studies involving different languages and scripts that the concept of a word - meaning conscious awareness of its function in both speech and writing - is dependent on knowledge of written language (Olson, 1996; Ramachandra & Karanth, 2007).

3.1.3 Syntax
The syntax of written language clearly differs greatly from speech. Spoken language is notoriously broken, choppy and incomplete. In contrast, the conventions of writing dictate a strict adherence to formal rules of grammar, including permissible constructions, punctuation and a preference for specific lexical alternatives (Chafe,
1985). This idealised written form serves as material for grammaticality judgements, which are the basis for understanding individuals’ knowledge about their own language. Although grammaticality judgements have served as primary data for developing theories of grammar, research on preliterate children and illiterate adults suggests that such judgements are based more on experience with written language than knowledge and use of spoken language alone.

Preliterate children judge sentence correctness based more on semantic than syntactic information, often ignoring word order violations (de Villiers & de Villiers, 1972). Children’s ability to judge a sentence based on grammatical information as opposed to the accuracy of its meaningful assertions does not appear until around the time they are learning to read and write (Olson & Nickerson, 1978; Hakes, 1980, cited in Olson, 1996). Similar behaviour has been observed in illiterate adults (Karanth et al., 1995; Karanth & Suchitra, 1983), leading researchers to conclude that the ability (inclination may be a better term) to analyse the structural properties of language separately from its content emerges only with literacy (Karanth & Sutchira, 1983).

3.2 Literacy and languages
Not only does literacy impact the language of individuals, but the use and dissemination of written material causes languages themselves to undergo dramatic change. Sometimes these changes are only reflected in the written form, but often the consequences of writing extend to spoken language as well.

Differences between spoken and written language include the use of more formulaic constructions in speech, colloquial vs. literary vocabulary and addition vs. subordination, among many others (Chafe, 1985). Writing follows the rules and conventions of idealised language much more strictly than spoken language (Linell, 1982). Although many of these may be flouted in everyday speech, the structure of writing can still influence that of spoken language. Written language is associated with the educated and elite, so it carries a high level of social prestige. It therefore comes to be perceived as the ‘real’ or ‘correct’ language, and social pressures can lead to it being used or required in speech (Linell, 1982).
The vocabularies of literate cultures are vastly expanded compared to preliterate cultures. Dictionaries and encyclopaedias act as storage for new, old and alternative words, which in turn enhances the precision of word meanings available to speakers (Linell, 1982). The preservation of words in print also retains meanings that are obsolete to present-day speakers. In preliterate cultures there is no way for unused terms to survive the temporal separation of speakers, so vocabularies consist of only those words used in everyday personal experiences (Ong, 1982).

Wray & Grace (2005) cite literacy as a cultural force that can profoundly affect linguistic form. The wide distribution and availability of texts means that readers are disconnected from the context of the author- culturally, spatially and temporally. To compensate for the lack of nonverbal strategies available in face-to-face communication, writers must use language that reliably and explicitly conveys complex meanings, ones that are often novel to the reader. This situation can induce regularisation of a language, as regular forms are both more likely to be applied to new words and understood by an unfamiliar reader. Lieberman & al. (2007) investigate a striking example of this tendency towards regularisation, that of the English regular past tense morpheme, of which the number of irregular forms has decreased dramatically since approximately 800 A.D. Along with other social and cultural factors, such as contact with Norse and French speakers and the global expansion of English through colonisation, the invention of the printing press and widespread literacy have surely contributed greatly to this pattern of regularisation.

3.3 Discussion
3.3.1 Summary
It is clear that literacy profoundly changes the way a person processes and thinks about language. Written language provides a new dimension of representation, one that includes information pertaining to pronunciation, syntax and meaning, and the composite of all three (Ehri, 1985). In other words, grammatical information is encoded in orthography in ways that are not apparent in speech. The accompanying metalinguistic awareness is precisely what may be influencing the participants of Cornish (2006) and Tamaraz & Smith (submitted), leading them to expect, discover and reproduce systematic languages. These studies present words in text, so both the modality of the stimuli and the prior knowledge of participants allow them to analyse
input using sophisticated and metalinguistically-informed strategies. Words can be analysed based on their phonemic/graphemic components into subunits, which the learner will expect to correspond with subunits of meanings.

The memory-enhancing effects of orthography also probably aid in the learning and reproduction of systematicity. If more words can be remembered during training, many examples of signal-meaning pairs will be available to form hypotheses about the structure of the language as a whole. A large inventory of examples enables the learner to analyse signals into their subunits, which is presumably taking place alongside the dissection of meanings into their components. The more examples a learner has of a particular signal unit being paired with a particular meaning unit, the more likely this correspondence will be noticed. Once a relationship between meaning and signal subunits is discovered, the learner will then expect the language to contain other consistent relationships and seek them out.

Orthography makes the compositionality of language explicit and tangible. The systematicity bias found by Tamariz & Smith (submitted) and evident in Cornish (2006) is thus very likely the result of participants utilising knowledge gained from literacy - not the operation of some innate feature of cognition. Another ALL study (Flaherty, 2007) examined iterated learning in children, who had much less experience with orthography than the adult participants of the other studies, and found no increase in structure, offering further support for this hypothesis. The experiment undertaken to investigate the influence of literacy on the systematicity bias will be discussed in Chapters Five and Six. However, remaining issues concerning literacy and linguistics warrant consideration.

3.3.2 Consequences for linguistics
The impact of literacy on both individual linguistic behaviour and larger-scale language change has important implications for the discipline of linguistics itself. Linell (1982) describes a ‘written language bias’ and asserts that linguists have traditionally been occupied with written language, often trivialising deeply-rooted differences between it and spoken language. Grammaticality judgements serve as primary linguistic data from which to develop linguistic theories, but we have seen
how those judgements are dramatically altered (and actually only possible) with literacy.

Grammatical structure and its correctness is in fact much less important for the interpretation of spoken language, which relies heavily on nonverbal strategies such as situational relevance and shared communicative context (Linell, 1982). Written language must use precise constructions that obey an extensive rule system because such strategies are completely absent from the medium of text. Linguists take this ideal, highly systematic form to be representative of typical linguistic material, and therefore end up with an inaccurate description of language. Linell (1982) cites this as cause for natural language being conceived of as a formal language. In addition, recursion, considered a central mechanism of syntax and possibly the only truly unique component of human language (Hauser, Chomsky & Fitch, 2002), may be merely an ‘artefact of analysis’ (Bickerton, 2008). Bickerton does not implicate literacy specifically, although it is arguably a primary cause of the intellectual biases responsible for this alleged misinterpretation.

A biased conception of language structure as a formal set of rules has consequences for the study of language in other contexts (Grace, 2002). The next Chapter will address how taking account of the qualitative differences between language as a spoken and written phenomenon relates to the study of language evolution.
Chapter 4  *Language Evolution*

4.1 Characterising language

4.1.1 Literacy and linguistic uniformitarianism

The study of the origins of language necessarily rests upon theories of language itself. In order to formulate an evolutionary explanation of how humans came to possess language, we must first have a clear and accurate description of the behaviour. Grace (2003) voices concern over the tendency of linguists to assume that the idealised form of language characterised by strict adherence to complex grammatical rules is representative of language in general. He notes that linguists’ failure to recognise the instrumental role of literacy and other cultural forces has led to an understanding of language ‘in which the products of more recent cultural evolution are lumped indiscriminately with the original products of our biological evolution’ (2003, pg. 1). Grace claims that erroneous theoretical assumptions can be attributed to the fact that linguists have traditionally described the language of texts, which we have seen differs greatly from speech (section 3.2). The language of literate cultures, as both a behaviour and a system of communication, has many features that are not shared by language as it exists in preliterate contexts.

Newmeyer (2002) extends this argument and questions the assumption that language has been the same since its inception\(^1\). Because it is traditionally seen as stemming directly from biology, it follows that language has not changed qualitatively since first used by the earliest humans (Bickerton, 1990; Pinker, 1994). Arguments have been made for the validity of an evolutionary scenario in which syntax is a product of biological natural selection (Pinker & Bloom, 1990). Such arguments are necessary only with the expectation that the earliest speakers used language as it is academically described today. However, if cultural processes influence language in significant ways, then it is possible that the starting point of language differed radically from its state today (Wray, 1998). Grammar is moulded by the needs of its users, so if language is used for more and very different purposes today (particularly in industrialised, literate societies), then it follows that grammar today is different from what it used to be (Newmeyer, 2002).

---

\(^1\) This line of reasoning ultimately derives from Lyell’s (1830) uniformitarian assumption in geology.
We do not, of course, have any direct way of knowing the nature of language and communication as it was early in human (or hominin) history. The task then becomes to understand precisely what language is as it occurs outside the influence of new, technology-enhanced environments. This knowledge, combined with what we know and will continue to learn about cultural processes affecting language, will allow us to come closer to understanding language as a more fundamental phenomenon, unelaborated by auxiliary forces. This is not to suggest that it is possible to know what language is in a vacuum; exogenous factors contribute to language development from the moment sensory information can be processed (DeCasper & Spence, 1987), so it is not useful or even coherent to ask what the nature of language is in the absence of culture or other environmental stimuli. The issue that must be resolved is to what extent our current understanding of language is in fact based on the influence of demonstrably modern cultural inventions, such as writing. Only then can we develop informed hypotheses as to how such a cognitive ability arose through evolutionary processes.

4.1.2 Syntax as a by-product of symbolic relationships

Not enough is known about the languages of preliterate societies, in part because so few of these exist, and also because much of the relevant research has been subject to the Western intellectual biases discussed previously. However, some authors have developed theories on the nature of language based on other available empirical evidence, and these have taken a perspective rather different from a rigid set of logical rules. Deacon (1997) claims that the most critical feature of human language is its symbolic nature. In Saussurean terms (Saussure, 1916), this encompasses both the relationships between the referent, signifier and signified, but also the intricate and elaborate conceptual links between referents. Deacon argues that the complex structure of human language and thought derives from the symbolic units we are able to form and the relationships between those; syntax is simply a by-product of semantics. In other words, systematicity in language naturally follows from the systematicity of our conceptual structures. Other authors (O’Grady, 2008; Schoenemann, 1999) also consider syntax an emergent property. Schoenemann (1999) questions its centrality as the distinguishing feature of human language and argues that grammatical structure is a cultural convention - converged upon through cultural processes - used to convey accurately to others multi-layered meanings.
constructed from semantic units\(^4\). This emphasis on semantics suggests that the traditional concept of language may include extraneous features, ones that are not specifically or directly related to our biological linguistic abilities.

It should also be noted that systematic analytic constructions are not always necessary to convey highly structured conceptual relationships. Formulaic or idiomatic expressions can equally be used to convey complex meanings and are ubiquitous in language (Wray, 2002). Depending on the social forces operating on a particular linguistic form, the structure of conceptual relationships may eventually come to be reflected in language structure through cultural-historical processes. The fact that holistic items remain a major feature of language, particularly speech, despite the usefulness of analytic constructions further supports the view of syntax as a consequence of transmission and communicative pressures, not a direct product of biological mechanisms (i.e. genes).

4.1.3 Child language acquisition

Another area of research that can shed light on the nature of language is child language acquisition. Determining what the acquisition process entails will reveal what the language faculty actually is. We have already seen that researchers are discovering the importance of culture in acquisition (section 2.1), and careful studies have been undertaken to try to pinpoint the independent contributions of biology and environment. ALL experiments offer a means to discover what children are able to learn and the mechanisms employed when doing so. The evidence suggests that children utilise general learning mechanisms for parsing linguistic data from a very early age (Maye et al., 2002; Gomez & Gerken, 2000).

Peters (1985) discusses how children utilise rudimentary devices to extract and segment meaningful aspects of speech, and this knowledge bootstraps further learning. The acquisition process is a gradual building of knowledge based on earlier learning. The more knowledge that is acquired, the more sophisticated are the strategies available to analyse linguistic input. This position holds that children come equipped with very general innate learning abilities. Later linguistic progress is not a

\(^4\) The term *units* should not be interpreted as being atomic; meanings in this case can be extraordinarily complex, encompassing the ideas, emotions and information content of, say, a story or argument.
product of separate innate language abilities, but from new abilities resulting from prior learning. In other words, experience shapes learning itself, so even if the initial mechanisms used by children are simple, the knowledge acquired with them changes how information is perceived and processed, and opens up new possibilities for learning. This account explains why literacy would have such a profound effect on learning in ALL tasks. The knowledge gained from literacy has altered participants’ cognitive mechanisms for parsing linguistic input, which in turn influences the strategies available to them for learning.

4.1.4 Ape language research
Research carried out by Savage-Rumbaugh and colleagues (Savage-Rumbaugh & Lewin, 1994) provides support for this alternative view of language. Despite previous, mostly unsuccessful, attempts to intentionally teach various ape species language (Patterson & Cohn, 1990; Terrace, 1979; Gardner & Gardner, 1969), bonobos under the care of Savage-Rumbaugh’s research team came to learn and use language to an unprecedented degree. The bonobos can understand spoken English, and their own utterances exhibit syntactic regularities (Greenfield & Savage-Rumbaugh, 1990). Remarkably, they did not undergo explicit instruction, but instead learned naturalistically, just like human children. Simply by experiencing language as it is used to communicate in a human environment, the apes came to possess abilities and exhibit behaviours that differ radically from anything that occurs spontaneously in the wild.

Savage-Rumbaugh (1998) documents a curious trend in the interpretation errors of one bonobo, Kanzi. Kanzi could accurately differentiate between syntactically-dependent, grammatically complex sentences, though he consistently performed poorly on seemingly simpler sentences that merely included extra objects, often failing to include one of the two items when acting out requests. Savage-Rumbaugh

---

5 This argument does not necessarily conflict with the proposal that children learn language holistically and are only later able to discover syntactic relationships (Wray, 1998). The pattern analysing skills described are apparently quite basic, probably shared by nonhuman primates (Tomassello, 2003a), and first applied to speech streams for phoneme and word learning (Peters, 1985). These may have the potential to detect complex grammatical structure, but that doesn’t mean to do so was their original function. That nonhuman primates can also learn fairly complex grammars (Fitch & Hauser, 2004) indicates that rudimentary pattern analysis skills can be utilised for learning that is not part of a species-typical cognitive or behavioural repertoire. Thus, children may preferentially or initially learn holistically, but because complex structure is now present in language due to cultural evolution, those same pattern analysing skills can be applied to language on a syntactic level.
speculates a deficiency in short-term memory, not grammatical competence, causes his observed errors. Somewhat counter-intuitively, grammatical complexity is not a problem for Kanzi as much as extra arguments to remember. This anomalous behaviour may involve how a string is analysed into its syntactic constituents (Truswell, 2008), which could be related to verb-argument structure and how Kanzi conceptualises objects and actions, indicating a deficiency in semantics, not syntax.

Taylor (Savage-Rumbaugh et al., 1998) argues that cognitive development in humans and (at least) bonobos is critically dependent on the social environment in which they are raised. He claims that pattern analysing skills possessed by both species become devoted to what is socially significant in the environment, and early commitment of these devices to a complex, highly structured behaviour such as language changes the course of development. The bonobos demonstrate the principle put forth by Peters (1985) regarding human language acquisition: future learning ability and potential intelligence is based not only on innate learning mechanisms, but on how experience interacts with these, altering learning and behaviour in profound and unexpected ways.

The stark differences between the bonobos and humans become most apparent in the realm of semantics (Savage-Rumbaugh et al., 1998). Their vocabularies only number in the few hundred, and their concepts are limited to the concrete, with even moderately abstract concepts causing them difficulty. What the bonobos can and choose to communicate about is incomparable to the metaphor, fiction and layers of abstraction that humans use. These deficiencies are consistent with the idea that systematicity naturally follows from the structure of conceptual space (Schoenemann, 1999; Deacon, 1997). The bonobos’ utterances are not as systematically complex as humans’, but neither are their concepts.

This research demonstrates that the ability to parse highly systematic relationships can be acquired with ‘primitive’ (or simply different) mechanisms given appropriate experience. The bonobos developed this ability without the biological ‘language faculty’ presumably possessed by humans, indicating their syntactic competence is solely and indisputably a product of learning. The fact that humans are equipped with similar learning abilities and develop similar systematic behaviours suggests that
human syntactic competence results from learning and experience, rendering obsolete the need for innate linguistic knowledge to account for this aspect of language. What remains is the rich and layered conceptual world manifested in human communication, which can more plausibly be attributed to biological differences between the two species.

4.1.5 Summary
The evidence presented suggests that the most powerful component of language is conceptual structure (discussed in detail by Hurford, 2007), embodied in semantics and the lexicon, and, purportedly, manifested in syntax. It then follows that the view of syntax as a necessary and fundamental feature of the biological language faculty is not correct. Systematicity in language emerges through language use over many generations of speakers (Schoenemann, 1999), and can become pronounced under certain socio-cultural pressures (Wray & Grace, 2005). Humans can learn to comprehend and produce highly systematic meaning to signal mappings, but this appears to depend on experience, not an innate, syntax-specific mechanism. What is innate would seem to be a capacity to form complex symbolic relationships, along with the ability to recognise and analyse complex patterns. This latter skill, however, is primitive and based on general learning mechanisms, as bonobos are able to accomplish a similar feat.

Developing an accurate and complete description of language, with the individual roles of biology and culture properly delineated, can inform theories on how it evolved. The following section will review the major positions on the issue and their compatibility with the research presented.

4.2 Language origins
Traditional perspectives on language evolution focus on biological processes and how natural selection could lead to language, in many cases specifically addressing features of phonology, semantics and syntax (Jackendoff, 2002; Nowak & Komarova, 2001; Bickerton, 1990; Pinker & Bloom, 1990). However, as the research reviewed suggests, much of language is likely not rooted in biology, and therefore does not require a biological evolutionary explanation. Moreover, features deriving from biology, such as pattern analysing skills, are components of general cognition, not
language specific. These mechanisms were therefore not under adaptive pressure for linguistic purposes, at least not solely, and thus natural selection for specific linguistic skills need not be invoked to explain their presence.

A number of other authors have offered alternative perspectives, emphasising cultural evolution and the interaction between culture and biology (Kirby & Christiansen, 2003; Smith, 2003; Tomasello, 2003a; Kirby & Hurford, 2002; Hurford, 1998; Deacon, 1997). Language is considered an adaptive entity in itself, which responds to the selective pressures induced by human cognition and communication (Christiansen & Ellefson, 2002). Hurford (2007) has argued that language should not be viewed as an unprecedented, monolithic trait that evolved independently in the hominid lineage, but instead as a ‘mosaic’, one that has co-opted subsystems present in many other species and merely quantitatively enhanced their functions. The combination of these two views - language as a cultural entity and biologically continuous in evolutionary terms - results in a much more parsimonious evolutionary scenario than one that explains many linguistic features through adaptive biological evolution.

Schoenemann (1999) argues that this characterisation is much more consistent with biological evolutionary principles. He claims that increasing semantic complexity is the major component of language that was under direct selective pressures, with syntax emerging later through communication of these conceptual structures. This position is certainly consonant with the evidence presented regarding systematicity, which suggests it is a predominantly learned trait, and thus does not require an evolved biological component entirely unique to humans.

Cultural evolution can ultimately influence biological evolution (Feldman & Laland, 1996), and Deacon (1997) has proposed a scenario in which the brain and language co-evolved. As systematicity became more prominent in communication systems due to cultural forces, there was perhaps some selective pressure for the ability to learn structurally complex systems with greater ease. Even so, this does not mean that highly systematic behaviours would become innate, merely the capacity to acquire increasingly complex systems would be sufficient. The latter requirement could simply enhance or build upon existing cognitive abilities, such as pattern analysis, a
possibility in line with Hurford’s (2007) proposal of evolutionary continuity and supported by ape language research (Savage-Rumbaugh et al., 1998).

4.3 Discussion
Developing accurate theories of language is a necessary prerequisite for investigating its origins. Research on literacy, recent work in linguistics, and ape language studies all indicate that syntactic structure is an emergent consequence of transmission pressures, not an innately pre-specified formal rule system. However, a cultural origin for syntax remains mostly speculative, and empirical verification is still required. Determining that systematicity in natural languages is an acquired trait - the result of learning how to learn a particular type of structured complexity - would have profound implications for theories of language evolution. Demonstrating that cultural processes are primarily responsible for linguistic structure would substantially reduce the explanatory burden of biological evolution. Theorising on the origins of language is then rendered a much more tractable task. The cultural processes implicated in language evolution remain in operation and take place on a much shorter timescale, and are thus observable to a much greater degree than biological ones.

Literacy research suggests that the strong systematicity bias reported by Tamariz & Smith (submitted) is a consequence of enculturation in a written system. If the bias is determined to be a product of learning, it will demonstrate that culture and experience can dramatically alter linguistic behaviour, and therefore, ultimately, language structure. By extension, perspectives on language origins emphasising the significance of cultural processes will receive support.

The ideal design to test for the effects of literacy in ALL tasks would be to compare illiterate and literate participants’ performances in learning an aurally-presented language, as orthography influences linguistic processing even in the absence of text (Ehri, 1985). However, a number of issues are associated with this paradigm. Finding adults who cannot read or write in an urban, Western European setting, particularly without accompanying learning or attention disorders, is not feasible. It would be impossible to find literate participants matched for characteristics such as cultural background and education levels. Moreover, because literacy in a society can eventually change the language itself (Linell, 1982), the linguistic behaviour of
illiterates will ultimately be affected. Attempting to contact and undertake research with a preliterate society is beyond the scope of the current project. Instead, an alternative method must be devised to circumvent the complications inherent to orthographic literacy. Chapter Five will describe one approach to observing how knowledge of a modern, cultural system affects systematicity in a similar task.
Chapter 5  *An Alternative Approach*

In many ways language appears remarkably unique among cognitive capacities; however, music bears many noticeable similarities (Jackendoff, 1980; Swain, 1997). Although, critically, music does not convey propositional meaning (Swain, 1997), a number of striking parallels with language exist. Music has a syntax of its own based on sequential patterns and combinations of notes, which has been described in the same terms as linguistic theories of grammar (Lerdahl & Jackendoff, 1983). Just as languages utilise different syntactical devices, the particular rules for musical composition are not universal and depend on cultural conventions (Sloboda, 1985). Musical syntax, moreover, is processed in similar areas of the brain to linguistic syntax (Patel et al., 1998). In addition, musical styles and composition, like languages, are affected by the forces of cultural evolution (Temperley, 2004). Particularly relevant for current purposes, a notational system is used for transcribing music.

There are at least some primitive universals involved in perceiving and processing music, such as the notion of a tonal scale (Sloboda, 1985). As with language, though, much of musical processing is culture-specific and learned. For instance, the precise categorisation of musical elements into discrete units on a scale is different across cultures (Sloboda, 1985). Common musical skill, such as humming a familiar tune or identifying an incorrect note, develops without any explicit instruction through the everyday social use of music (Sloboda, 1985). Some skills reach surprising levels of sophistication, and nonmusicians perform similarly to musicians on many musical processing tasks (Bigand & Poulin-Charronnat, 2006). However, musical training alters processing dramatically in many respects. Trained musicians develop an awareness of structural units and features. As a result, their internal representation of music becomes more intricately structured and memory capacity for musical sequences is increased (Sloboda, 1985). Sloboda notes that although nonmusicians have implicit knowledge of musical structures, training provides a tool to describe and think about the same structures.
Ehri (1985) directly compares written language to musical notation, which adds a visual-spatial level of representation to musical knowledge. In the same way that all language users are proficient speakers and listeners, nonmusicians are ‘experienced listeners’ of music, able to recognise and reproduce familiar melodies (Bigand & Poulin-Charronnat). Musical training and notation spurs the development of metamusical awareness not typically possessed by the casual music consumer, just as metalinguistic knowledge resulting from literacy is not found in preliterate children or illiterate adults.

The cognitive and structural parallels between music and language provide a more feasible approach to investigating the relationship between literacy and the systematicity bias. Moving to a musical modality should still allow us to examine how explicit knowledge gained from acquiring a written system to represent normally implicit knowledge changes behaviour related to that knowledge. We assume musical literacy has similar cognitive consequences for musical awareness and processing as orthographic literacy does for language. If we suspect that orthographic literacy causes the observed systematicity bias in ALL tasks, we can hypothesise that musical literacy will induce a similar effect in a related musical task. Indeed, the fact that critical differences related to conveying meaning do exist between music and language means that positive results in a musical paradigm will provide strong evidence that access to abstract representations is the primary factor responsible for the systematic behaviour of participants in the reviewed ALL studies.

5.1 Musical language learning

A musical paradigm must be developed with the aim of mirroring as closely as possible the factors associated with orthography and language. We expect learning a musical language to be extremely difficult for both nonmusicians and musicians, as even musical training does not involve the coupling of referential content to musical sequences. Traditional ALL tasks are themselves demanding, so transferring the stimuli to an aurally-presented, unfamiliar modality will likely increase the effort required from all participants; therefore, considerations must be made to ensure that learning is still possible. It is also important that the features of a musical language are analogous to the phoneme/grapheme relationship between speech and
orthography, so the units comprising signals must be represented in musical notation. In addition, distinctive signal units are needed to replace the syllable inventories used in the original studies. The character of these units must be salient enough for both groups to perceive, and simple enough to be reproduced with relative accuracy, at least preserving the distinctive relationships between signals. One way to fulfil these conditions is to use intervals on the Western musical scale.

5.1.1 Components of a musical language
Intervals are readily perceived by nonmusicians and musicians alike (Smith et al., 1994)⁶. They are also particularly important distinctions for musicians, being the basis for harmonics and other aspects of musical grammar. Experienced musicians can identify intervals reliably and accurately (Sloboda, 1985). Intervals are represented in musical notation, so we can expect musicians to have an internal, visual-spatial representation of these relationships. So, even though both groups will perceive distinctions made in the signal space, musicians will be equipped with abstract representations to attach to signal units. Based on self-reports of participants in other ALL studies (Tamariz & Smith, submitted), we can expect all participants to look for patterns between meanings and signals. However, musicians’ meta-musical knowledge should aid them both in discovering which interval goes with which meaning component, and also in remembering corresponding pairs (assuming some compositionality). If musicians can recognise the structural units of a system, they may be more likely to realise that a language follows combinatorial rules (if it does) or impose their own system for attaching signals to meanings.

Perceptibility and discriminability of intervals must be considered, as those extending beyond an octave will be difficult to distinguish and identify. In addition, participants will be required to reproduce intervals vocally (see section 6.1.1 for reasons behind the choice of output form), so frequency range is also a factor. For these reasons, intervals based from middle C, C4 (262 Hz), are ideal (see Figure 1). Each signal unit will consist of a pair of tones, the first of which is always C4. Moving up the

⁶ It is clear that nonmusicians can perceive intervals; identification of these is where a divergence in ability between musicians and nonmusicians arises. However, Smith et al (1994) found that nonmusicians' performance in an identification task is similar to musicians when nonmusicians are given a token based on a segment of a familiar tune to label intervals. Once their knowledge is tapped with the appropriate tools, nonmusicians can apparently access their own abstract representations of music, although it appears that explicit instruction is necessary for this to occur.
scale, the second note can be $E_4$ (330 Hz), $G_4$ (392 Hz), $B_4$ (494 Hz) or $C_5$ (525 Hz). Unison ($C_4$) can also occur as the second note. Moving down in pitch, possible notes include $B_3$ (247 Hz), $G_3$ (196 Hz), $E_3$ (165 Hz) and $C_3$ (131 Hz).

This system was adopted because even without explicit knowledge of intervals, listeners can perceive that a pair of notes will either be the same note repeated, or the second note will be higher or lower to some varying degree. For pairs with less extreme pitch distances, participants, particularly nonmusicians, might have difficulty in identifying whether the notes are higher or lower; however, this should not impede reproduction necessarily, as people are able to casually hum or sing tunes without this type of knowledge. Moreover, because of the extreme variation in high and low pitch combinations, nonmusicians should still be able to infer a system based simply on higher or lower note combinations. So, even if underlying structure remains outside conscious awareness, they can still produce their own similar system accordingly.

![Figure 1](image_url) Intervals used as signal units to construct tunes. Symbols along the base represent notes, and numbers along the lines above denote the distance in semitones between $C_4$, the base note for every pair of tones, and the other notes.

5.1.2 Measuring learning
We expect very few musicians and nonmusicians to reproduce the intervals precisely; it is unlikely that participants from either group possess the vocal control to do so. The nature of intervals means that participants can adjust their output to a key within their normal frequency range, something that naturally occurs when repeating familiar tunes (Sloboda, 1985). Even so, participants will probably not reproduce the intervals with great accuracy. However, they can still produce some system of intervals
resembling their input. Because learning the language exactly is highly improbable, we can instead focus on whether participants have learned the system of a language. If participants can distinguish between signal variants and realise that there is some regular mapping between them and meaning variants, they can attempt to replicate these same distinctions in their output, even if those end up being different from the precise form of their input. Using these criteria, we can determine if participants are recognising and applying systematicity, regardless of whether their actual production matches the input. This method relieves some of the factors that may handicap nonmusicians (see section 6.1.1 for further discussion of these issues), and thus means differences in overall output structure found between the two groups can be more safely attributed to sensitivity to systematicity in the input.

It will still be useful to measure the similarity in form between inputs and outputs. Although we do not expect output to closely resemble input, examining the extent to which it does or does not will tell us how well a participant has learned exact form in addition to the relationships between meaning-signal pairs. There are limiting factors in determining this measure. For one, there is a divide between reception and production. Even if the precise signal for a meaning is known, the likelihood that participants will produce it perfectly is very low. Moreover, if an interval of -2 is used where -1 occurred in the input, the nature of the musical scale does not allow us to say that -2 is more similar to -1 than, say, -4. The perceptual salience of musical intervals is not straightforward or linear (Russo & Thompson, 2005a; 2005b). For this reason, a participant will only be given credit for reproducing a form if it is the exact interval as in the input. We expect musicians to do this more than nonmusicians, as they should recognise that the signal set is based on familiar intervals. However, how this measure relates to other measures of output structure, input structure and the relationship between the two can elucidate the processes underlying learning in both groups.
It should be noted that some musicians may possess relative pitch\(^7\), which would of course increase the accuracy of form reproduction. However, by simultaneously measuring reproduction of form and system we can see how this ability interacts with systematic behaviour.

### 5.2 Summary

Music is a universal human behaviour (Sloboda, 1985), has structural properties resembling the syntax of language (Lerdhal & Jackendoff, 1983), and has a culturally developed, visual-spatial system of representation. The parallels between music and language allow us to adapt traditional ALL studies by developing an artificial musical language to investigate the nature of the systematicity bias discussed above. The next Chapter will discuss in detail a study utilising such a musical language that mirrors an orthographic literate-illiterate comparison.

---

\(^7\) Relative pitch is defined as the ability to identify the musical interval separating two simultaneously or successively played notes, or the ability to produce a named interval above or below a given note (Sloboda, 1985). Absolute pitch, or ‘the ability to name or produce a note of a given pitch in the absence of a reference note’ (Deutsch, 2006; pg. 11) is not an issue in this case, as all stimuli consist of sequences of notes.
Chapter 6 The Present Experiment

The following sections describe a musically-based experiment adapted from the ALL studies reviewed in section 2.2.3.

6.1 Methods
A miniature musical signal set was paired with meanings to produce a set of musical languages (see Figure 2 for a breakdown of the meaning space). A meaning-signal pair consists of an image paired with a short tune. Tunes corresponding to a composite meaning are composed of three pairs of notes (intervals). In the fully compositional language, each interval corresponds to an individual meaning component. Colour is encoded in the first pair, shape in the second and insert in the third. The partially compositional language follows a similar pattern, but with exceptions for each meaning component. The random language has no regular mapping between signals and meanings and can be considered entirely holistic. (see section 6.1.3 for details of the structure of each input language.)

As we are most interested in the processes involved in individual learning, no generational transmission was adopted. Three separate languages were created with varying levels of compositionality (none, partial and full), and separate groups of musicians and nonmusicians were trained on each language. A 50% bottleneck was imposed during training, but testing required participants to provide responses for every meaning. This was done for two reasons. First, results from pilots indicated that learning a musical language is difficult to the point of becoming overwhelming. If participants were trained on all 27 meaning-signal pairs, the potential frustration could hinder learning. Second, we are not simply interested in the ability to learn a language, but instead how participants respond to varying levels of systematicity – whether musicians will discover, learn, reproduce and possibly create systematicity. A bottleneck requires participants to create signals of their own for meanings not previously encountered based on partial knowledge of the language, thus enabling innovation. We expect this process will be implicit and participants mostly unaware of novel meanings during testing, as they will have experience with all meaning components from training.
Overall learning and reproduction is expected to be substantially reduced compared to traditional ALL studies, though we still anticipate musicians’ performance to proportionately mirror that of literates in orthographic tasks. The opportunity for innovation provided by the bottleneck allows us to determine if participants have discovered underlying patterns, when they exist, and generalised those to novel items.

6.1.1 Counteracting unrelated advantages from musical training
Musicians naturally appear to be at an advantage in learning a musical language, but steps were taken in order to best address the issue at hand. As mentioned, the bottleneck inhibits participants from simply memorising musical sequences and necessitates innovation. If the systematicity bias is in operation with musicians, we can expect their output – a full inventory of meaning-signal pairs – to be more systematic overall, regardless of whether they are reproducing their input exactly. The same situation gives nonmusicians the opportunity to impose patterns on the language, even if they cannot remember specific sequences as well as musicians.

As another step taken to avoid indirect advantages of musical training, participants were required to hum or sing their output. If a keyboard or some other device was used to produce tunes, musicians’ experience with instruments (of any type) would put them at an advantage for navigating perceptual space. Both groups will have at least limited experience with humming or singing. Of course, tone deafness and the inability to repeat a tune with any degree of accuracy would hinder performance, but these can be detected through a simple screening procedure.
Participants from both groups will probably not learn their input languages to a great extent at all, creating many of their own signals when prompted with meanings during testing. Exactly how musicians versus nonmusicians handle this opportunity to innovate can tell us a great deal about how the input is perceived and represented by each group. First generation participants of Cornish (2006) only successfully learned input languages that contained a sufficient level of structure, though learnability in this case was based on similarity measurements. Although we expect to see a similar increase in learnability for the more compositional languages, we anticipate that what will be better learned is the structure of the languages, not the languages themselves. Therefore, we can instead measure and describe the system of mapping from meanings to signals used by participants to determine how well they have learned their input. So, the nature of the task may change what we are measuring, but the pattern of measurements for musicians should still mirror the original results.

6.1.2 Method of Analysis

Structure was quantified using RegMap, a measure of the regularity of meaning-signal mappings based on information theory and weighted by the frequency with which items in a system are encountered (Tamariz & Smith, submitted). RegMap differs from other methods of quantifying structure (Kirby et al., 2008; Cornish, 2006; Tamariz, 2008; Shillcock et al., 2001) in that it is not based on similarity. Instead, it calculates co-occurrence frequencies between units of signals and meanings, and so is independent of any perceptual similarities in either dimension. RegMap essentially measures to what extent a system of meanings and signals departs from randomness. It can describe systematicity on multiple levels- for a single signal or meaning variant, a single signal or meaning dimension (referred to as systematicities in the analysis) and also for a language as a whole (referred to as compositionality). RegMap values range from 0 (random) to 1 (fully compositional).

RegMap is sensitive to frequency, and therefore to a bottleneck. It determines how much a learner can expect a particular meaning to map to a particular signal based on how many times this pairing occurs. Thus, if a language has a compositionality level of 1 as a whole language, when a 50% bottleneck is accounted for, the compositionality level will actually be lower. This is because a learner cannot be
fully certain that a meaning will pair with a given signal, as each instance has not been encountered.

*Regmap* is particularly useful for analysing and describing the data from this experiment, as we are more interested in the learning of system than form (section 5.1.2). It can pinpoint precisely where structure is manifested in a language, so values can be compared between languages to see not only how overall structure compares, but also internal structure. Similarity measurements are often used to determine how well a language has been learned (Kirby et al., 2008; Cornish, 2006; Flaherty, 2007). However, constraints on production in this task will render similarity measurements inaccurate, causing differences between input and output to appear inflated. We can avoid this problem by instead measuring how well participants learn the system of relationships between meaning-signal pairs in a language. For these reasons, *RegMap* is the ideal method of measurement.

### 6.1.3 Materials

The experiment was administered using a *Psycscope* script adapted from Tamariz & Smith (submitted) and run on an Apple Macbook. Signals were replaced with sound files. Individual tones were created in Adobe Audition by generating pure tones 330 ms in duration. Tone pairs, which constitute a single signal unit, were created by inserting 150 ms of silence between a *C4* and each of the designated nine tones. Tunes were created by inserting 550 ms of silence between pairs.

The unstructured language, L1, was created by randomly combining possible signal units into strings of three, which were then randomly assigned to meanings. This was done until a sufficiently unstructured language was produced, with a compositionality level of 0.00458 for the whole language and 0.00395 with the bottleneck. L3, the fully compositional language, was constructed by randomly assigning each signal unit to one meaning dimension, and then creating whole tunes by combining these signal units systematically (see section 6.1). The compositionality level of L3 as a whole is 1, becoming 0.42546 with the bottleneck. L2 was created by systematically altering individual meaning-signal subunits until its compositionality level reached 0.66615 as a whole and 0.29390 with the bottleneck. Items to be included in the bottleneck were randomly chosen for L1 and L3, but multiple randomly assigned bottlenecks were
generated for L2 until the ratio of its compositionality levels was proportionate to that of the full and bottleneck conditions for L3 (see Appendix I for all versions).

Tunes were played through headphones, and participants’ responses were recorded via a headphone-mounted microphone onto a Microtrak recorder. The frequency analyser was at a resolution of 65,536 samples/second. A programme employing the RegMap method was used to quantify the systematicity of languages. Systematicities are calculated using the following formula (from Tamariz & Smith, submitted):

\[
Systematicity = \frac{\Sigma(RF_i)}{\Sigma(F_i)} \times \frac{1}{d}
\]

Where \( R \) is the redundancy of a signal dimension variant across all possible meaning variants, which indicates the certainty of a signal variant encoding a particular meaning variant. \( F \) is the frequency of a signal dimension variant in the language, and \( d \) is the discrepancy between the number of signal and meaning variants, defined as the greater divided by the lesser of the two. The compositionality level of languages is measured by substituting signal and meaning variants in the equation with signal and meaning dimensions (systematicities). Another program was used to generate random languages from the signal inventories in participants’ output (see section 6.2.1).9

6.1.4 Procedure
Runs were conducted individually and took approximately 30 to 45 minutes. Participants were told they would be learning a musical language and trained on a series of images paired with tunes. They were asked to practise repeating tunes aloud each time one was played during the training phases. In addition, they were told to hum or sing a tune for every image they saw in the testing phase, even if they were unsure about the correct tune. Finally, prior to beginning the experiment, participants were played and required to repeat two six-note tone sequences10, and their responses recorded. This was done to test their ability to repeat accurately a tone sequence of the same length as those in the language. If a participant was unable to do so, the

---

8 Redundancy is a mathematical concept defined as the opposite of entropy for a limited set of mutually independent variables (Shannon, 1948)
9 The code for both programs was provided by Monica Tamariz.
10 These sequences were of a similar temporal structure as those used in the languages, but the notes and combinations were not the same.
experiment continued, although their data was subject to closer inspection to
determine if it should be included in the final analysis. Training and testing
procedures for each participant were as follows:

1. Instructions
2. Training
   2.1 Round 1
      2.1.2 Image displayed and tune played (tune repeats three times with 6 seconds
           of silence in between for participant to vocally repeat) X 14
   2.2 Break
   2.3 Round 2 (same as Round 1)
   2.4 Break
   2.5 Round 3 (same as Round 1)
3. Break and instructions for testing phase
4. Testing
   4.1 Image displayed and participant asked to provide correct tune X 27 images
5. End

Responses were recorded during both the training and testing phases. The
experimenter was in the same room for the pretest. However, to prevent inhibition
about providing vocal responses, the experimenter left after the pretest and moved to a
small room just off the main room, from which it was still possible to hear and
monitor participants. When the computer-based task was complete, the experimenter
returned to the main room and asked participants a short series of questions to obtain
more detailed information on their knowledge of the language and strategies for
learning it (Appendix II). Each participant received £5 as compensation.

Output was analysed using Adobe Audition 1.5, which assigned the frequency of each
note in a sequence to the appropriate semitone. These were then used to calculate the
intervals for each signal unit. Due to the nature of the software, some value
judgements were necessary. If the semitone assigned by the frequency analyser was
obviously incorrect, a keyboard was used to determine independently the closest
semitone. In addition, the accuracy of the software was intermittently verified
throughout the analysis of each participant’s output by playing the assigned semitones
on a keyboard and checking these against the voice recording. In some instances,
particularly when a participant attempted to reach a very high or very low note, the
volume or timbre of the voice precluded any frequency analysis. In these cases, the

If a participant performed poorly on the pretest, their practise responses during the training phase
were examined in more detail to determine if they could consistently repeat a sequence with more
practise. If their performance improved during training and responses became consistent, their data
was kept, otherwise, it was discarded.
note was assigned to whatever the lowest semitone the participant had successfully produced. If no semitone approached such an extreme, the note was arbitrarily assigned to a semitone at the high or low end of a participant’s range, usually the octave. This strategy resulted in what is probably a fairly accurate analysis, as participants would regularly fail to produce a very high or very low note within their output, so the system was internally consistent. Finally, because very few participants produced perfect semitones, sometimes the combination of actual pitches produced resulted in potentially ambiguous intervals. For instance, the first note in a signal unit might be analysed as C4+49 and the second as G4-4912. In this case, the interval might actually be smaller than seven semitones, as both notes are bordering the next highest and lowest semitones, respectively. When this occurred, the three possible semitone combinations (C4-G4, C#4-G4 or C4-F#4) were played on a keyboard and checked against the voice recording. The interval that most closely resembled the actual output was assigned to the signal unit.13

6.2 Results and discussion
A total of 37 participants were enlisted. Due to equipment malfunction (1), tone deafness (2) and procedural error (4), only data from 30 were included in the final analysis. Musicians (M) totalled seven females and ten males, mean age 24.3 years. Nonmusicians (NM) totalled nine females and four males, mean age 24.1 years. Requirements for inclusion in the M group were the ability to read musical notation proficiently, play an instrument proficiently and to be currently practising. The mean length of time practising was 17.5 years. NM were required to have no musical background. No participants had experience in linguistics beyond an undergraduate introductory course. In addition, no participants spoke a tonal language, as experience with tone and meaning might lead to biased processing of a musical language.

Participants were asked to rate the difficulty of the task on a scale of one to five, with one being easy and five being very difficult. The expectation that both groups would consider the task difficult was largely confirmed, with mean ratings across all languages being 4.73 (NM) and 4.15 (M). Ratings, on average, did not decrease for

12 These numbers represent cents, or hundredths of a semitone.
13 This was generally according to the opinion of the experimenter, although in particularly difficult cases, a disinterested third party was asked for judgement. In addition, output was analysed blind; meanings that corresponded to participants’ tunes were not available during auditory analysis.
more compositional languages. The duration of the experiment was virtually the same across groups and conditions (Ms - mean: 28.5 minutes; NMs - mean: 28.2 minutes).

Comparing the output of musicians and nonmusicians can be approached in multiple ways, each one commenting on a somewhat different aspect of the results. First, we can compare compositionality levels of individual languages, both for participants of the same input groups as well as between input groups. Second, we can look at how both groups responded to the structure of their input languages. Finally, we can examine how faithfully participants reproduced the form of their input.

6.2.1 Compositionality of individual languages

The way participants go about mapping meanings to signals reveals not only how well they learned the input, but also how their own output is structured, independently of how the input may have been. Table 1 shows participants’ compositionality levels for groups and languages. Levels for Ms are, on average, higher than those of NMs, with this differentiation becoming more marked as the structure of the input increases. A Two-Way Independent ANOVA showed this relationship to be nonsignificant for the main effect of music ($p = .148$) but closer to significance for the main effect of language ($p = .087$). Tests also indicated a nonsignificant interaction effect between music and language ($p = .579$). Planned contrasts revealed the effect of language was significant between L1 and L3 ($p < .05$), but not between L1 and L2 ($p = .250$) or L2 and L3 ($p = .226$). Contrary to test results, the nonparallel lines in Figure 3 suggest a significant interaction effect (Field, 2005). The visible interaction between group and language type could indicate that the significance between L1 and L3 stems primarily from musicians (see discussion).
<table>
<thead>
<tr>
<th>Language</th>
<th>Musicians</th>
<th>Nonmusicians</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language 1</td>
<td>0.00395</td>
<td>0.00169</td>
</tr>
<tr>
<td></td>
<td>0.00182</td>
<td>0.00500</td>
</tr>
<tr>
<td></td>
<td>0.00063</td>
<td>0.00162</td>
</tr>
<tr>
<td></td>
<td>0.00615</td>
<td>0.00053</td>
</tr>
<tr>
<td></td>
<td>0.00615</td>
<td><strong>0.00221</strong></td>
</tr>
<tr>
<td></td>
<td>0.00262</td>
<td>0.000348</td>
</tr>
<tr>
<td>Language 2</td>
<td>0.01412</td>
<td>0.00468</td>
</tr>
<tr>
<td></td>
<td>0.00717</td>
<td>0.00133</td>
</tr>
<tr>
<td></td>
<td>0.00387</td>
<td>0.00631</td>
</tr>
<tr>
<td></td>
<td>0.00083</td>
<td>0.00135</td>
</tr>
<tr>
<td></td>
<td>0.05129</td>
<td>0.00425</td>
</tr>
<tr>
<td></td>
<td>0.00061</td>
<td><strong>0.00358</strong></td>
</tr>
<tr>
<td>Language 3</td>
<td>0.03625</td>
<td>0.00661</td>
</tr>
<tr>
<td></td>
<td>0.00234</td>
<td>0.00238</td>
</tr>
<tr>
<td></td>
<td>0.07576</td>
<td>0.00251</td>
</tr>
<tr>
<td></td>
<td>0.42212</td>
<td>0.00211</td>
</tr>
<tr>
<td></td>
<td>0.00912</td>
<td><strong>0.00340</strong></td>
</tr>
<tr>
<td></td>
<td>0.01315</td>
<td><strong>0.09312</strong></td>
</tr>
</tbody>
</table>

**Table 1** Compositionality levels of individual participants’ output from both groups for each condition. Numbers in bold are means.

**Figure 3** Chart of means for musicians and nonmusicians for each language, showing a clear interaction between group and language type. (Note that because the data was transformed in order to meet assumptions of normality, lower values indicate higher compositionality.)
We can also examine individual meaning-signal dimension pairs (systematicities) within participants’ outputs. The systematicity for each meaning dimension can be calculated to determine how much a particular meaning is encoded in a particular signal unit. Structure in a language is dependent upon the number of variants in the signal inventory. Since this is not constant across output languages, it may not be a reliable method of analysis to compare directly systematicities between them. Converting systematicities to p-values, however, allows us to compare directly the significance of structure for a given meaning-signal dimension pair.

P-values can be calculated by generating many random languages from the meaning and signal variant inventory of each participant and measuring systematicities for each (Monte Carlo analysis, $N = 1,000$). The corresponding participant’s systematicities can then be compared to these, and the ranking of a participant’s systematicity represents the p-value for that pair. Table 2 shows p-values for each participant and each meaning-signal dimension pair.

In L1, very few p-values reached significance; however, Ms had three significant and one highly significant value, while NMs had only one significant value. For L2, four of the highly significant values for Ms occurred in the pairs that were systematic in their input, while only one of these did for NMs. In L3, every one of Ms’ highly significant values (plus two significant ones) occurred in the systematic pairs of their input, while NMs had only one instance of this, although it was not highly significant. The pattern of scores in L2 and L3 indicates not only that Ms’ languages are more structured than NMs’, but also that this structure more closely resembles that of their input.
<table>
<thead>
<tr>
<th>Language 1</th>
<th>Color S1</th>
<th>Color S2</th>
<th>Color S3</th>
<th>Shape S1</th>
<th>Shape S2</th>
<th>Shape S3</th>
<th>Insert S1</th>
<th>Insert S2</th>
<th>Insert S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>0.006</td>
<td>0.117</td>
<td>0.223</td>
<td>0.420</td>
<td>0.154</td>
<td>0.364</td>
<td>0.223</td>
<td>0.061</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>0.301</td>
<td>0.304</td>
<td>0.020</td>
<td>0.700</td>
<td>0.727</td>
<td>0.922</td>
<td>0.563</td>
<td>0.585</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>0.333</td>
<td>0.000</td>
<td>0.938</td>
<td>0.604</td>
<td>0.640</td>
<td>0.588</td>
<td>0.221</td>
<td>0.411</td>
<td>0.257</td>
</tr>
<tr>
<td></td>
<td>0.137</td>
<td>0.703</td>
<td>0.900</td>
<td>0.555</td>
<td>0.415</td>
<td>0.981</td>
<td>0.952</td>
<td>0.045</td>
<td>0.451</td>
</tr>
<tr>
<td></td>
<td>0.702</td>
<td>0.198</td>
<td>0.798</td>
<td>0.817</td>
<td>0.916</td>
<td>0.498</td>
<td>0.093</td>
<td>0.002</td>
<td>0.807</td>
</tr>
<tr>
<td>NM</td>
<td>0.311</td>
<td>0.934</td>
<td>0.802</td>
<td>0.933</td>
<td>0.339</td>
<td>0.395</td>
<td>0.417</td>
<td>0.133</td>
<td>0.331</td>
</tr>
<tr>
<td></td>
<td>0.302</td>
<td>0.062</td>
<td>0.209</td>
<td>0.470</td>
<td>0.766</td>
<td>0.633</td>
<td>0.132</td>
<td>0.183</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>0.239</td>
<td>0.067</td>
<td>0.317</td>
<td>0.887</td>
<td>0.777</td>
<td>0.318</td>
<td>0.817</td>
<td>0.460</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td>0.442</td>
<td>0.811</td>
<td>0.055</td>
<td>0.232</td>
<td>0.386</td>
<td>0.182</td>
<td>0.787</td>
<td>0.268</td>
<td>0.049</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language 2</th>
<th>Color S1</th>
<th>Color S2</th>
<th>Color S3</th>
<th>Shape S1</th>
<th>Shape S2</th>
<th>Shape S3</th>
<th>Insert S1</th>
<th>Insert S2</th>
<th>Insert S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>0.200</td>
<td>0.569</td>
<td>0.692</td>
<td>0.598</td>
<td>0.580</td>
<td>0.694</td>
<td>0.436</td>
<td>0.001</td>
<td>0.417</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.498</td>
<td>0.096</td>
<td>0.728</td>
<td>0.388</td>
<td>0.007</td>
<td>0.842</td>
<td>0.935</td>
<td>0.764</td>
</tr>
<tr>
<td></td>
<td>0.491</td>
<td>0.772</td>
<td>0.939</td>
<td>0.774</td>
<td>0.518</td>
<td>0.443</td>
<td>0.844</td>
<td>0.564</td>
<td>0.747</td>
</tr>
<tr>
<td></td>
<td>0.118</td>
<td>0.639</td>
<td>0.442</td>
<td>0.064</td>
<td>0.393</td>
<td>0.139</td>
<td>0.785</td>
<td>0.555</td>
<td>0.324</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.073</td>
<td>0.593</td>
<td>0.883</td>
<td>0.001</td>
<td>0.590</td>
<td>0.893</td>
<td>0.956</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.134</td>
<td>0.520</td>
<td>0.351</td>
<td>0.341</td>
<td>0.281</td>
<td>0.383</td>
<td>0.631</td>
<td>0.870</td>
<td>0.435</td>
</tr>
<tr>
<td>NM</td>
<td>0.359</td>
<td>0.887</td>
<td>0.705</td>
<td>0.609</td>
<td>0.315</td>
<td>0.809</td>
<td>0.369</td>
<td>0.460</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>0.610</td>
<td>0.582</td>
<td>0.469</td>
<td>0.044</td>
<td>0.823</td>
<td>0.960</td>
<td>0.049</td>
<td>0.821</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>0.017</td>
<td>0.280</td>
<td>0.005</td>
<td>0.196</td>
<td>0.739</td>
<td>0.113</td>
<td>0.302</td>
<td>0.433</td>
<td>0.616</td>
</tr>
<tr>
<td></td>
<td>0.533</td>
<td>0.057</td>
<td>0.313</td>
<td>0.936</td>
<td>0.736</td>
<td>0.510</td>
<td>0.691</td>
<td>0.935</td>
<td>0.623</td>
</tr>
<tr>
<td></td>
<td>0.254</td>
<td>0.475</td>
<td>0.925</td>
<td>0.286</td>
<td>0.541</td>
<td>0.350</td>
<td>0.110</td>
<td>0.802</td>
<td>0.112</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language 3</th>
<th>Color S1</th>
<th>Color S2</th>
<th>Color S3</th>
<th>Shape S1</th>
<th>Shape S2</th>
<th>Shape S3</th>
<th>Insert S1</th>
<th>Insert S2</th>
<th>Insert S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>0.041</td>
<td>0.675</td>
<td>0.933</td>
<td>0.787</td>
<td>0.001</td>
<td>1.002</td>
<td>0.777</td>
<td>0.650</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.072</td>
<td>0.973</td>
<td>0.968</td>
<td>0.968</td>
<td>0.215</td>
<td>0.023</td>
<td>0.319</td>
<td>0.033</td>
<td>0.492</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.783</td>
<td>0.589</td>
<td>0.874</td>
<td>0.001</td>
<td>0.900</td>
<td>0.963</td>
<td>0.818</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.793</td>
<td>0.661</td>
<td>0.970</td>
<td>0.001</td>
<td>0.994</td>
<td>0.988</td>
<td>0.997</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.474</td>
<td>0.362</td>
<td>0.723</td>
<td>0.009</td>
<td>0.001</td>
<td>0.722</td>
<td>0.250</td>
<td>0.816</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>0.140</td>
<td>0.813</td>
<td>0.789</td>
<td>0.650</td>
<td>0.957</td>
<td>0.062</td>
<td>0.602</td>
<td>0.271</td>
<td>0.001</td>
</tr>
<tr>
<td>NM</td>
<td>0.073</td>
<td>0.187</td>
<td>0.035</td>
<td>0.004</td>
<td>0.168</td>
<td>0.395</td>
<td>0.855</td>
<td>0.060</td>
<td>0.673</td>
</tr>
<tr>
<td></td>
<td>0.320</td>
<td>0.116</td>
<td>0.290</td>
<td>0.007</td>
<td>0.140</td>
<td>0.598</td>
<td>0.555</td>
<td>0.164</td>
<td>0.481</td>
</tr>
<tr>
<td></td>
<td>0.666</td>
<td>0.661</td>
<td>0.022</td>
<td>0.490</td>
<td>0.681</td>
<td>0.854</td>
<td>0.173</td>
<td>0.056</td>
<td>0.838</td>
</tr>
<tr>
<td></td>
<td>0.911</td>
<td>0.101</td>
<td>0.469</td>
<td>0.410</td>
<td>0.424</td>
<td>0.772</td>
<td>0.100</td>
<td>0.184</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Table 2 The systematicities for each meaning-signal dimension pair for each participant converted into p-values. Numbers highlighted in blue indicate significance below .05, those highlighted in yellow indicate significance below .005. Columns bordered in bold lines are those in which meaning-signal mapping was systematic in L2 and L3.

6.2.1.1 Discussion
The ANOVA showed the effect of music to be nonsignificant for compositionality levels between languages; however, close inspection of values reveals that individual variation within musicians is quite high. For L1, variance is very low for both groups (M = 6.46E-06; NM = 3.75E-06). Within musicians it increases for both L2 and L3 (0.0004 and 0.0267). The same does not occur with nonmusicians, whose variation remains low for both languages (4.78E-06 and 4.59E-06). Within both L2 and L3, a few musicians produced remarkably regular languages. We can infer that they have
learned a system - not simply memorised their input - because their regular patterns extend to novel items only encountered in testing. Although their compositionality levels appear quite low, the requirement of vocally producing output likely obscures the systematic mapping that a participant may have intended to employ. Even so, Table 2 shows that M#5 in L2 mapped from meanings to signals in a highly significant systematic way that mirrored the input. Table 1 shows that compositionality levels are quite high for three musicians in L3 (#s 1, 3 and 4), with one reaching the same level as the input (#4). No nonmusician attained compositionality levels anywhere close to these musicians’ partially compositional languages.

Looking at L2 and L3 in the p-value analysis, it is interesting to note in what dimension pairs regularity is manifested for musicians versus nonmusicians. Not only do those of musicians more closely mirror the input, but there is further evidence to suggest that nonmusicians are not processing input in the same way. NM#3 in L2 has two instances of significant systematicity, although these are both for color. The participant is encoding color in a more holistic manner, with the beginning and ending of a sequence signifying this meaning dimension. The structure of sequences has not been analysed compositionally, in contrast to musicians, who, in the vast majority of cases, encode each meaning dimension in a separate signal dimension.

There is some fundamental difference in the way the two groups are processing the languages, indicated by the fact that some musicians were clearly able to learn and reproduce the compositional system of their input while no nonmusicians performed at a similar level. The nonsignificant results of the ANOVA are probably a consequence of small sample sizes. If more participants were enlisted, the distinction between musicians and nonmusicians would likely be much more pronounced. High individual variation within musicians brought means down in L2 and L3, which may be another reason for nonsignificant statistics. Individual variation can be high in traditional ALL tasks (Cornish, 2006; Tamariz & Smith, 2008), and it appears to be even higher in the present experiment. Although musicians might have greater meta-musical awareness than nonmusicians, even they are not accustomed to pairing referential content to musical sequences. Specifically expecting phonemes and/or graphemes to correspond to meaning is part of the metalinguistic awareness brought
on by orthographic literacy (Ehri, 1985; Nunes et al., 2006), so the fact that the bias for systematicity is evident in more participants in a textual modality is not surprising. Nevertheless, the fact that some musicians produced highly regular languages demonstrates that musical literacy can also provide a tool for analysing input and discovering systematic relationships.

6.2.2 Structure of output compared to input
For this method of analysis, the weighted ratio of compositionality levels for output to input was calculated by multiplying the ratio of output to input by the sum of the input and output. This process makes comparing all three languages more useful, as otherwise even a small increase in compositionality for learners of L1 would result in a very high ratio. By incorporating information from input languages, this figure can tell us how input influences the level of structure in a learner’s output.

For L1, on average, the ratio for Ms was slightly higher than NMs (0.0078 vs. 0.0042). For L2 and L3, ratios for Ms were, on average, higher than NMs (L2: mean 0.0147 vs. 0.0036; L3: mean 0.1658 vs. 0.0034). A Two-Way Independent ANOVA showed the main effect of musical training to be nonsignificant ($p = .149$), the main effect of language nonsignificant ($p = .323$) and the interaction effect nonsignificant ($p = .594$). Despite these negative results, a closer look at the data reveals a clear interaction between group and language, as can be seen in Figure 4. Again, the extremely nonparallel lines representing L2 and L3 are suggestive of a significant interaction.14 The ratios for Ms increased markedly between L2 and L3, while those of NMs stayed about the same (see discussion).

Systematicities for each meaning-signal dimension for both input and output were calculated to determine if a correlation exists between the two (this is the same as the ‘learnability’ measured in Tamariz & Smith, 2008). This measure shows where structure is manifested in a language, and comparing the two will indicate if the internal structure of an output language mirrors that of the input. Because L1 lacked internal structure, and a correlation between input and output would not indicate that participants had learned a system of mappings, analysis will be restricted to L2 and

14 The crossing lines for L1 and L2 are also suggestive of an interaction effect; however, this interaction is more likely due to the nature of the analysis and the very low compositionality of L1.
L3. For L2, systematicities were found to be significantly correlated for one M (p < .01; r = .88) but no NMs. For L3, there was a significant correlation for one M (p < .05; r = .60), highly significant correlations for three Ms (p < .01 for all; r = .94, .88 and .96) and no significant correlations for NMs. The correlation coefficients were then compared to the compositionality levels of the corresponding participants, and these were found to be significantly correlated for Ms (p < .05; r = .56) but not NMs (p > .10, r = -.16).

![Weighted ratio for output/input](chart.png)

**Figure 4** Chart displays the means for each group and language type. Crossing lines indicate an interaction effect. (Note that data were transformed in order to meet normality assumptions, so lower figures indicate higher ratios).

6.2.2.1 Discussion

Musicians appear to be more sensitive to the structure of their input than nonmusicians. As in the previous measures, small sample sizes and individual variation may be a major cause for nonsignificant ANOVA results. Even so, the performance of some musicians suggests their response to systematic input differs markedly from nonmusicians. Although the compositionality levels of nonmusicians increased marginally from L1 to L2, the compositionality of L3 does not seem to aid them in discovering, reproducing and extending this system as it does for musicians.
Two musicians increased the systematicity of L1, while only one nonmusician did so. Although this may suggest musicians are imposing systematicity on the language, the increases in compositionality are very slight. It is probably the case that even if musicians were able to dissect the signal and meaning space, they were unable to detect any patterns, and thus attempted to learn the language holistically. The marked difference in ratios between groups is most apparent in L3. The ratios become lower for nonmusicians because the compositionality level of the input increased from L2 to L3 while that of their output stayed about the same as L2. However, for musicians the increase in the input did not cause ratios to decrease because the outputs also increased. In fact, ratios became larger, indicating a sensitivity to the higher regularity of the input.

A correlation between meaning-signal dimensions is an important measure of learning, particularly for the present experiment. If systematicities of an output language correlate with the input, the system of mapping from meanings to signals has been successfully learned. So, even if the compositionality level of participants’ output remains relatively low (due to production limitations, etc.), it is still possible to know if they have recognised and attempted to replicate the regular mappings found in the input. Correlations were strong for musicians in L2 and L3, while no significant relationship was present for nonmusicians. In addition, musicians with higher compositionality levels also had higher correlations in this measure, while the relationship was actually negative for nonmusicians. This final correlation indicates that the nonmusicians who produced more structured languages did so by employing a very different system from their input. Therefore, we can conclude that they were not aware of the structure of the input to the level that musicians were, and so did not analyse input in the same way.

6.2.3 Similarity of form
Examining if and how well participants learned the precise form of their input provides additional information on how the two groups went about learning the languages. The presence of a bottleneck also makes this measurement useful. If participants who reproduce the form of a signal unit generalise and apply it
systematically to novel meanings, we can infer that they have analysed the system of a language successfully.

We measure similarity of form by counting the number of signal units in a participant’s output that were identical to the signal unit for the same meaning in the input. We distinguish between units encountered in training (‘seen’) and units first encountered during testing (‘unseen’, but correctly generalised). Ms averaged 11.3 seen units and 6.6 unseen; NMs averaged 3.6 seen and 1.4 unseen. A Two-Way Independent MANOVA showed the main effect of musical training to be significant (Pillai’s Trace, \( p < .05 \)), the main effect of language to be nonsignificant (\( p = .206 \)) and the interaction effect to be nonsignificant (\( p = .411 \)). The effect of musical training was significant for seen items (\( p < .01 \)) and unseen items (\( p < .05 \)). Language was significant for unseen items (\( p < .05 \)) but nonsignificant, though nearing significance, for seen items (\( p < .07 \)). Planned contrasts showed that language became significant for both variables between L1 and L3 (seen: \( p < .05 \); unseen: \( p < .05 \)). Although the test returned a nonsignificant interaction effect, Figures 5 and 6 show that an interaction is clearly present between groups and language types. Again, the extremely nonparallel lines are suggestive of a significant interaction (see discussion).

![Image](image-url)

**Figure 5** An interaction is clearly visible from the nonparallel lines representing language type.
6.2.3.1 Discussion
Musicians faithfully reproduced significantly more seen items than nonmusicians, as well as used the correct form on significantly more unseen items ($p < .05$). Compositionality level of the input also had a significant effect, especially on unseen items. At first glance, it would seem that musicians are simply better at memorising and reproducing musical form; however, the interaction graphs illustrate an important factor in their ability to do so. Both groups perform virtually identically in L1, with neither correctly reproducing their input language to a great extent. This means that neither group is learning L1, as the only reliable measure of this would be similar form. Only when the compositionality of input increases do musicians begin to improve, with both seen and unseen items rising dramatically from L1 to L2 and L2 to L3. Nonmusicians’ performance increases little over all languages for both measures. Compositionality levels of musicians’ output are significantly correlated with the number of seen items they reproduce ($p < .001$, $r = .88$), while this is not the case for nonmusicians ($p > .10$, $r = .33$). It would seem musicians are only able to learn the form of a language correctly when they have also discovered the underlying structure.
Musicians who learned and reproduced the form of seen items also extended the pattern to novel meanings. This indicates they are not simply memorising the input, but that they have analysed the signal and meaning structure of the language correctly. In fact, on some occasions, musicians trained on L2 actually corrected some instances of irregularity in their input, applying the regular pattern to these items in their output. No such correction occurred in nonmusicians.

It is also interesting to note how participants reacted to novel meanings. Most musicians (particularly those who performed well on all measures) claimed to recognise all images in testing even though nearly 50% were novel, while some nonmusicians reported not recognising many of the images. These musicians had learned the structure of the language to a greater degree, and thus analysed meanings into their constituent parts, so perhaps novel composite meanings did not look unfamiliar because they are simply a recombination of familiar components. That nonmusicians tended to notice new images suggests their concept of the meaning space was, like the signal space, holistic.

Musical literacy, like orthographic literacy, appears to create an awareness of subunits, which are the indicators of structure in a system. Awareness of subunits allows a learner to analyse input in a way that reveals structure, which explains the starkly different performances of musicians and nonmusicians on form measures. The marked increase in faithful reproduction of more compositional languages by musicians mirrors the increase in learnability for the literate participants of Tamariz & Smith (submitted) and Cornish (2006), further solidifying the dependence of the systematicity bias on analytical tools provided by cultural systems.
Chapter 7 General discussion

7.1 Summary of results
Each method of analysis has demonstrated that musicians are learning and processing the musical languages in an altogether different manner than nonmusicians. Statistics showed group differences to be nonsignificant, though other factors, such as the combination of high individual variation and small sample sizes, likely weakened the impact of any significant effects. Close examination of the data reveals that interesting trends are still evident. Despite negative test results, there are still strong indications that musical literacy will induce a systematicity bias in a musical language learning task.

The pattern of systematicity in musicians’ output closely resembles that of literates in the original ALL studies. Those trained on a random language produce mostly unsystematic languages. Structure increases somewhat for learners of a partially compositional language. This increase does not seem to be as marked in the present experiment, although the presence of a bottleneck may be responsible for the discrepancy. The absence of 50% of the language in training means that regularities were not reinforced over irregularities to the same extent as for participants in the original task. The reduction of regular examples may have contributed to the observed reduction in systematicity of output. Finally, for the fully compositional input, output structure increases markedly for some musicians.

In contrast, nonmusicians’ output structure does not steadily rise, remaining at similar levels for all three language types. It can be noted that nonmusicians’ compositionality levels did increase, however marginally, from L1 to L3, though this should not be interpreted as contradicting our hypothesis. Nonmusicians were highly literate individuals who already possess linguistic and metalinguistic skills, which were probably enlisted while attempting to learn the musical languages. However, it is clear these skills alone were not sufficient to discover and generalize systematicity to anywhere near the same level as some musicians.
The structure of musicians’ highly systematic outputs mirrored the structure of the input. Moreover, the more systematic a musician’s output, the more its structure resembled the input. Nonmusicians did not follow a similar trend, with the few somewhat structured outputs employing a system that differed greatly from the input.

Reproduction of form appears to be related to reproduction of system. Musicians’ performance on similarity measures was virtually identical to nonmusicians’ for the random language, and only surpassed them when input compositionality levels were high. This suggests the ability to learn and reproduce precise form is dependent on the structure of a language. The meta-musical awareness that aids in detecting structure probably includes meta-knowledge of the signal units (intervals). Discovery of the underlying pattern involves recognising what comprises the signal inventory. Therefore, knowing the structure of the system implies knowing its components. We can further conclude that musicians are not simply better at memorising input than nonmusicians, as they apply the correct form to novel meanings. Ultimately, precise reproduction depends heavily on awareness of structure in addition to skilled vocal control.

Although some musicians do not reproduce the form or structure of their input, those that do illustrate clearly the effect that musical training and literacy can have on learning a musical language. No musician learned the random language to anywhere near the same degree as the partially and fully compositional languages. The musicians who successfully learned these languages are responding to different levels of compositionality, preferentially learning more systematic languages and exhibiting similar behaviour to literates in orthographic tasks. There were no instances whatsoever of highly-accurate learning in nonmusicians. Therefore, we can reason that the systematicity bias observed in some musicians can be attributed to musical training and literacy.

7.2 Meta-musical awareness and evidence for the systematicity bias
The observed reduction in the number of participants exhibiting a systematicity bias may in part be because meta-musical awareness does not involve the same expectation for pairing form and referential meaning as metalinguistic awareness. Although musical structure can convey emotional content, it is not conventionally
used to convey informational content (Swain, 1997). This may explain why fewer musicians detected language structure. In addition, knowledge of musical notation, the major criterion for participation in this study, may not be the best measure for meta-musical awareness. Many talented musicians never learn to read music, yet their knowledge of musical structure is surely extensive. Abstract representations of music are available from other means, such as finger positions on guitar frets, which offer another type of visual-spatial representation perhaps more salient than symbols on a page. These abstract representations, though not textual, are still analogous to those provided by literacy. If another measure of meta-musical awareness could be developed and used to select participants for a similar study, clearer evidence for a systematicity bias may be found.

7.3 Broader implications

It was hypothesised that the systematicity bias observed in the highly literate participants of earlier, text-based ALL studies (Tamariz & Smith, submitted; Cornish, 2006) was the result of experience with written language. The fact that a systematicity bias is also evident in a music-based task is quite remarkable, given the novelty of the paradigm and lack of a propositional meaning component encoded in musical structures. This indicates the bias largely results from meta-knowledge of the structural features of a symbolic system; learning and reproduction of systematicity are aided by abstract representations of structure. The findings strongly support the hypothesis that the bias observed in the original studies stems from orthographic literacy.

The results strongly suggest that the systematicity bias is predominantly a learned trait, not the work of innate mechanisms. This would remove the conflict in timescales between cultural explanations of language structure (Kirby, 2000; Wray & Grace, 2005) and data from the original ALL studies. These results can serve to inform future empirical research examining the learning and cultural transmission of language in human users. Studies must be undertaken with factors such as literacy in mind in order to ensure experimental designs and the data obtained accurately reflect the actual phenomena researchers aim to understand.
In conjunction with research on literacy, the findings illustrate the significant effect of experience on systematic behaviour, supporting arguments for the importance of culture and learning over innate knowledge in the acquisition of syntax. By extension, emergence theories of grammatical structure emphasising the centrality of semantics and symbolic relationships are also supported. Taken together, these issues have important implications for work in language evolution, stressing the instrumental role of cultural evolution in determining many features of language.
Chapter 8 Conclusion

The experimental results outlined above demonstrate that experience with a cultural system for explicitly representing normally implicit knowledge induces systematic behaviour. This outcome supports the hypothesis that the systematicity bias observed in the presented ALL studies is a consequence of orthographic literacy, reconciling those findings with accounts of cultural processes gradually causing language structure to reflect conceptual structure. Syntactic structure, particularly compositionality, therefore may well be a matter of culture and learning, not innate properties of cognition. The biological or cultural status of syntactic abilities is of course not fully resolved, and the acquisition process in children must be carefully studied to understand the development of syntactic competence more completely. However, these results strongly support theories of language that emphasise culture and learning as dominant forces affecting individual linguistic behaviour and language structure.

This work also highlights the critical need to consider literacy when undertaking linguistic research. Linguists developing theories of grammar must incorporate relevant research from other disciplines, ensuring their theories accurately describe primary characteristics of human language. The field of evolutionary linguistics is particularly sensitive to necessary adjustments in language theories, as evolutionary explanations directly depend on exactly what we understand language to be. A long-standing curiosity concerning the origins of language has produced much speculation, both amateur and academic, on how language might have evolved. The serious study of language evolution, however, remains in relative infancy, and significant progress cannot be made without first guaranteeing that the phenomenon under investigation is the correct one.
Appendix I Input Languages

Intervals and tunes for languages one, two and three. Bold numbers indicate a tune was included in the learning bottleneck (or ‘seen’ items).

### Language One

<table>
<thead>
<tr>
<th>Blue</th>
<th>Yellow</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>-12</td>
<td>-5</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>-12</td>
<td>11</td>
<td>-8</td>
</tr>
<tr>
<td>11</td>
<td>-4</td>
<td>7</td>
</tr>
<tr>
<td>-1</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>-5</td>
</tr>
<tr>
<td>-12</td>
<td>4</td>
<td>-5</td>
</tr>
<tr>
<td>-12</td>
<td>-8</td>
<td>11</td>
</tr>
<tr>
<td>-5</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

### Language Two

<table>
<thead>
<tr>
<th>Blue</th>
<th>Yellow</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>-12</td>
<td>-5</td>
<td>-1</td>
</tr>
<tr>
<td>-0</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>-8</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>-12</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>-8</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>-8</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>11</td>
<td>-1</td>
</tr>
</tbody>
</table>

### Language Three

<table>
<thead>
<tr>
<th>Blue</th>
<th>Yellow</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>-8</td>
<td>-5</td>
<td>-1</td>
</tr>
<tr>
<td>-8</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>-8</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>-8</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>-8</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>-8</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>-8</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>-8</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>-8</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Appendix II  Questionnaire

1. How difficult would you rate the task of learning the musical language (on a scale of 1-5)?

2. How did you go about trying to learn the language?

3. Did you notice any patterns in the tunes and meanings?

4. How confident were you that you knew the correct tune for each meaning?

5. Did you recognise all the meanings/pictures you saw during the testing phase?

6. If you couldn’t remember the exact tunes during testing, did you try to use a pattern of some kind?
References


Tamariz & Smith (submitted). Quantifying language regularity: Exploring the interactions between systematicity and learnability using a miniature artificial language experiment.


