LUND UNIVERSITY

Knowledge Representation, Heuristics, and Awareness in Artificial Grammar Learning

Johansson, Tobias

2008

Link to publication

Citation for published version (APA):
Johansson, T. (2008). *Knowledge Representation, Heuristics, and Awareness in Artificial Grammar Learning*. [Doctoral Thesis (compilation), Department of Psychology]. Department of Psychology, Lund University.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 18. May. 2025

Knowledge Representation, Heuristics, and Awareness in Artificial Grammar Learning

Tobias Johansson
Department of Psychology
2008

LUND UNIVERSITY

Printed in Lund, Sweden, 2008 by KFS in Lund AB
ISBN: 978-91-628-7563-3
© Tobias Johansson (2008)

Contents

Abstract

People can become sensitive to the general structure of different parts of the environment, often without studying that general structure directly, but through being incidentally exposed to instances that conform to the structure. When such learning proceeds unintentionally and gives rise to knowledge that is difficult to verbalize it is often referred to as *implicit learning*. One of the most commonly used experimental paradigms in the study of implicit learning is *artificial grammar learning*, in which participants are exposed to sequences that conform to a set of rules without being informed about the presence of rules. In a subsequent test phase, participants can usually distinguish between sequences that conform to and sequences that violate the rules, without being able to say much about the underlying rules. There are many different theories about the kind of knowledge representations that underlie sensitivity to general structure in artificial grammar learning, and there are also different viewpoints concerning how to measure the conscious status of the knowledge acquired in artificial grammar learning. Investigating these different theories is important, partly because it may provide an understanding of the extent to which complex learning and abstraction of structure proceeds unconsciously.

Study I of this thesis investigated artificial grammar learning and the use of a fluency heuristic, which involves relying on the surprising ease of processing an item as a basis for making a judgment. Other studies have shown that the fluency heuristic is used in a wide variety of judgments (e.g., recognition and preference). Study I showed that participants rely on a fluency heuristic in artificial grammar learning as well, but mainly under non-analytic processing conditions when participants were encouraged to respond rapidly and thereby make global judgments about items without processing details to any large extent. This is consistent with the idea that fluency may provide a cue for indirect sensitivity to general structure.

Study II investigated the effect of non-analytic processing on the conscious status of knowledge as assessed by confidence judgments. It was found that non-analytic processing increased the availability of conscious knowledge, consistent with the idea that part of the knowledge acquired in artificial grammar learning may be, not inherently unconscious, but of a kind that is available through a non-analytic form of introspection. One possibility is that, relative to more analytic forms of introspection, non-analytic introspection may be more sensitive to the non-focal peripheral contents of consciousness, the so called "fringe consciousness". This could explain why the knowledge acquired in artificial grammar learning often seems intuitive, even though it is not necessarily unconscious.

Study III investigated whether artificial grammar learning gives rise to knowledge that is independent from the surface features of the exposure material. A number of claims have been offered in the literature for such surface-independent knowledge, particularly as a result of extended exposure to regularities. The results clearly suggested that the knowledge formed under observational learning conditions in artificial grammar learning is not independent from the surface features of the exposure material. The results are consistent with a variety of computational models of artificial grammar learning that rely on surface-dependent perceptual representations.

Finally, Study IV investigated whether the knowledge acquired in artificial grammar

learning is unconscious in the sense that it may be expressed unintentionally. The results showed that, to the extent that knowledge was expressed, it was expressed intentionally. However, the low levels of performance in Study IV limit the generality of the findings. Possible reasons for the low performance are discussed in the context of different models of artificial grammar learning.

Taken together, the studies in this thesis illuminate issues regarding both knowledge representation and the conscious status of knowledge in artificial grammar learning. In general, the studies are in line with an episodic framework according to which the general abstract structure of a domain is not automatically extracted. Instead, both learning and awareness proceeds as a function of task demands, intentions, expectations, and processing strategies.

Studies

Study I

Johansson, T. (in press). In the fast lane toward structure in implicit learning: Non-analytic processing and fluency in artificial grammar learning. *European Journal of Cognitive Psychology*.

Preprint reproduced with permission from *European Journal of Cognitive Psychology*, © 2008 by Taylor & Francis. The article is available online at:
http://www.informaworld.com/smpp/content~content=a792977968~db=all~jumptype=rss

---

Study II

Johansson, T. (2008). Non-analytic processing and discrimination of knowledge states in artificial grammar learning. Unpublished manuscript, Department of Psychology, Lund University.

Study III

Johansson, T. (in press). Strengthening the case for stimulus-specificity in artificial grammar learning: No evidence for abstract representations with extended exposure. *Experimental Psychology*.

Preprint reproduced with permission from *Experimental Psychology*, Vol. 56 (2), 2009, © 2009 by Hogrefe & Huber Publishers · Cambridge, MA · Toronto · Göttingen · Bern.

Study IV

Johansson, T. (2008). Against the grain: Intentional control of structured behavior in implicit learning. Unpublished manuscript, Department of Psychology, Lund University.

Acknowledgements

Before going more straight to the point, I´d like to set the stage and provide some context and ramification for what´s about to come. Think of it as a backdrop against which to evaluate the nerve and presence of what I intend to communicate in this passage, but which I may not be able to put into words.

Great. Comfort. Achievement. Effort. Analyze. Produce. Expectation. Dark. Stuck. Back. Run. Hide. Panic. Silent. Scream. Sleep. New. Fresh. Start. Move. Fun. Complex. Lost. Back. Start. Basic. Walk. Explore. Failure. Back. New. Start. Old. Rail. Move. Hold. Move. Hold. Let go. Run. Panic. Breathe. Not so bad. Ok. Close. Look up. Take in. Content. Closure. Travel. Speak. Panic. Finish. Crawl. Up. Strike. Feedback. Crap. Perspective. Take in. Breathe. Take in. Breathe. Damn! Breathe. Take in. Calm down. Point of no return. Way back. Deadline. Closer. Deadline. Closer. End of the world. No. Still here. Alive. Think. Wonderful. Great. Comfort. Achievement. Trap! Down! Cover! Nope. False alarm. Where am I? Am I done?

That was the fast life version of my stroll down Ph.D. lane. Of course, I exaggerate, I am overly hindsight dramatic, and it´s been great fun along the way. But still, there´s a certain feeling of constant insecurity, looking over one´s shoulder, thinking “am I dead wrong?” and “is this really any good?”. It takes a while to get a grip on insecurities like that, and believe me, they flourish among doctoral students. One cure is understanding, another is having a life. I still work on the former, because you never really get there, and the latter could not be better. I consider myself fortunate to have received support in all these different ways from different people.

There´s a dual character to most things. Looking into my own head I still don´t know what to make of all this doctoral stuff. But as my dearly beloved keeps telling me, I´ll try to enjoy and appreciate having made it this far. It is an achievement, and I do take pride in it. Am I done? Well, yes. Done with *this*. No more doctoral insecurities… Let´s go get some new ones! Getting-a-job-and-career insecurities. They never end, do they? But neither does the pleasure that I derive from doing all this. Many of the reasons for that can be traced to other people. I´d like to be able to trace the patterns of the web of people responsible for making me enjoy all of this. But I am not much for naming names and making all those selections. I wouldn´t be able to paint an accurate picture of that anyway. You all probably know your part, and I´d rather let you know face to face. Still, I cannot finish this without a direct representation of what right now strikes me as having had the most profound impact on my doctoral efforts:

Bella – center of my universe and love of my life,
August – my son, dearest to me, pure love, smile, meaning,
Whitney – Lanchashire heeler, love, confidence, attitude,
Carl Martin – supervisor, open, smart, flexible.

Tobias Johansson

Summary in Swedish – Svensk sammanfattning

Knowledge Representation, Heuristics, and Awareness in Artificial Grammar Learning

Tobias Johansson
Institutionen för Psykologi
Lunds Universitet, 2008

Människor (och andra arter) har en grundläggande förmåga att anpassa sig till och lära sig hur omgivningen är strukturerad. Ofta lär vi oss struktur indirekt, utan att direkt och avsiktligt studera strukturen ifråga. Klassiska vardagsexempel på tämligen oavsiktlig strukturinlärning är språkinlärning, utveckling av musikaliska preferenser, inlärning av sociala regler, och, mer generellt, inlärning av olika typer av beteendemönster i förhållande till omgivningen. När inlärning av strukturella regelbundenheter sker oavsiktligt och ger upphov till kunskap som är svår att verbalisera så kallas det för *implicit inlärning*, vilket kan kontrasteras mot explicit inlärning som är avsiktlig och ger upphov till mer verbal kunskap.

En vanlig experimentell metod för att studera implicit inlärning är artificiell *grammatikinlärning*. I den här metoden får deltagarna observera sekvenser av symboler som följer ett underliggande komplext regelsystem, men deltagarna får inte till en början veta att sekvenserna följer regler. Efter att ha observerat sekvenser en viss tid så får deltagarna veta att de sekvenser de har sett följer regler och att de nu ska få observera nya sekvenser, varav hälften följer reglerna och hälften bryter mot reglerna. Uppgiften är nu att klassificera vilka sekvenser som följer reglerna. Ofta kan deltagarna klassificera korrekt bättre än slumpnivå, vilket indikerar att de har lärt sig något om den underliggande strukturen i sekvenserna, men deltagarna kan oftast inte säga särskilt mycket om reglerna som styr sekvenserna.

Det finns för närvarande många olika teorier om vilken typ av kunskap som underligger deltagarnas förmåga att klassificera strukturerade sekvenser i artificiell grammatikinlärning, och det finns också olika synsätt angående hur man bör mäta hur pass medveten deltagarnas kunskap är i de här sammanhangen. I den här avhandlingen har jag undersökt både kunskapsrepresentation och den medvetna statusen hos kunskap i samband med artificiell grammatikinlärning. Att undersöka dessa frågor utgör ett viktigt projekt, delvis eftersom det kan hjälpa till att belysa i vilken utsträckning komplex inlärning och abstraktion av struktur sker utan medvetna influenser.

Studie I i den här avhandlingen undersökte om deltagarna förlitar sig på ”fluency” när de klassificerar sekvenser i artificiell grammatikinlärning. Att förlita sig på fluency innebär att man tar beslut beroende på hur pass smidigt man bearbetar informationen ifråga. Testsekvenser (de man klassificerar i testfasen) som följer reglerna i artificiell grammatikinlärning delar oftast mer egenskaper med träningssekvenserna (de man såg i observationsfasen) än vad testsekvenser som inte följer reglerna gör. Detta möjliggör potentiellt högre fluency för testsekvenser som följer reglerna än de som inte följer reglerna. Studie 1 manipulerade fluency på ett artificiellt sätt i testfasen genom en teknik som kallas för maskerad priming och innebär att testsekvensen ibland visas väldigt snabbt på datorskärmen innan den återigen dyker upp på skärmen mer permanent för klassificering. Resultaten visade att deltagarna förlitar sig på fluency främst när klassificeringsbesluten togs under tidspress, vilket är i linje med att användandet av fluency ökas vid icke-analytiskt processande som betonar globala bedömningar av stimuli snarare än fokus på detaljer. Generellt sett visar resultaten att fluency kan utgöra en indirekt ledtråd till känslighet för struktur i sekvenser, men att användandet av fluency som en ledtråd inte är ovillkorligt.

Studie II undersökte effekten av icke-analytiskt processande i testfasen på den medvetna statusen hos deltagarnas klassificeringsbeslut. Ett vanligt sätt att mäta den medvetna statusen hos deltagarnas klassificeringsbeslut är att be deltagarna att avge konfidensbedömningar efter varje klassificeringsbeslut. Alltså, efter varje klassificeringsbeslut (”Den här testsekvensen följer/följer inte reglerna”) så får deltagarna säga hur säkra de är på att beslutet är korrekt (”Jag är *X %* säker på att beslutet är korrekt”). Om deltagarna är mer säkra på sina beslut när de är korrekta än när de är inkorrekta så tyder det på att deltagarna i viss mån vet om när de har rätt och när de har fel (dock inte nödvändigtvis *varför* de har rätt eller fel), vilket i sin tur är en betydelse i vilken kunskap kan vara medveten. Studie II visade att icke-analytiskt processande i testfasen kan öka graden av medveten kunskap, vilket är i linje med att en del av den kunskap som införskaffas i artificiell grammatikinlärning är, inte direkt omedveten, men av ett slag som är tillgänglig via icke-analytisk introspektion. En möjlighet är att icke-analytisk introspektion är mer känsligt för de mer perifera och subtila delarna av medvetandeinnehållet när deltagarna gör sina bedömningar. Detta kan i sin tur förklara varför deltagarna ofta har svårt att verbalisera det som de lär sig i artificiell grammatikinlärning och varför man ofta anser att kunskapen är tämligen intuitiv.

Studie III undersökte om artificiell grammatikinlärning ger upphov till kunskap som är abstrakt i bemärkelsen att kunskapen är oberoende av den perceptuella formen av inlärningsmaterialet. I en del tidigare studier har det hävdats att sådan abstrakt kunskap formas mer eller mindre automatiskt via implicit inlärning, särskilt i samband med förlängd inlärningsfas. Resultaten från Studie III visade tydligt att den kunskap som etableras via observationsinlärning i artificiell grammatikinlärning inte är oberoende av den perceptuella formen i materialet. Förlängd inlärningsfas leder till bättre kunskap om regelbundenheterna i sekvenserna, utan att samtidigt leda till mer abstrakt kunskap. Resultaten är i linje med diverse matematiska beräkningsmodeller av artificiell grammatikinlärning, nämligen modeller som utvecklar representationer som är knutna till den perceptuella formen av inlärningsmaterialet. Andra studier visar att abstraktion över perceptuella former kan ske i viss mån i artificiell grammatikinlärning, men sådan abstraktion verkar ske tämligen avsiktligt och medvetet.

Studie IV undersökte om artificiell grammatikinlärning ger upphov till kunskap som är omedveten i bemärkelsen att kunskapen yttrar sig utan att deltagarna har för avsikt att tillämpa kunskapen. Resultaten visade att, i den utsträckning kunskap om regelbundenheter tillämpades, så tillämpades kunskapen avsiktligt, vilket är i linje med att artificiell grammatikinlärning inte ger upphov till omedveten kunskap som är bortom deltagarnas kontroll. Det bör dock nämnas att denna slutsats begränsas av deltagarnas låga prestation i Studie IV. Olika möjligheter till den låga prestationen diskuteras i Studie IV i samband med olika modeller av artificiell grammatikinlärning.

Sammantaget så belyser de olika studierna i den här avhandlingen frågor gällande både kunskapsrepresentation och den medvetna statusen hos kunskap i samband med artificiell grammatikinlärning. Generellt sett så är resultaten i linje med ett episodiskt synsätt på artificiell grammatikinlärning. Enligt det synsättet så sker inlärning av struktur inte automatiskt, utan är istället en kombinerad funktion av de (direkta eller indirekta) krav som uppgiften involverar, deltagarnas intentioner, förväntningar, och de strategier deltagarna använder för att bearbeta informationen ifråga. Ett sådant synsätt har potential att föra samman resultat från olika domäner (t ex minnesforskning och kategoriseringsforskning) inom ett mer enhetligt perspektiv.

Knowledge Representation, Heuristics, and Awareness in Artificial Grammar Learning

Introduction

One of the most fundamental topics in psychology, arguably *the* most fundamental topic, concerns how knowledge is represented by the human mind. This thesis is about one particular aspect of that topic, namely the representation of knowledge that is acquired through implicit learning and that provides a basis for sensitivity to structural regularities.

Under standard and sober circumstances, people do not generally behave and think randomly. Instead, people act in structured ways, with varying degrees of complexity, efficiency, and accuracy. Most of our everyday behaviours depend on some form of sensitivity to structural properties of the world. Classic examples of such behaviours are language production (Perruchet, 2005), music appreciation (Kuhn & Dienes, 2005), and the execution of structured motor responses (Nissen & Bullemer, 1987).

In addition, it seems that learning of structural regularities often proceeds more or less unintentionally, without direct intention to learn the regularities in question. For example, learning a language partly involves becoming sensitive to various regularities of the relevant units of language. However, when children learn a language it is not as if they actively try to deduce or intentionally test their way through the infinite mud of possible descriptions of the regularities. Rather, the process seems to be largely unintentional, and the resulting knowledge can be put to use much more effectively than its content can be verbalized. Learning that proceeds unintentionally with respect to the structural properties in question and gives rise to knowledge that is difficult to verbalize is often referred to as *implicit learning* (Cleeremans, Destrebecqz, & Boyer, 1998; Shanks, 2005), a term coined by A.S. Reber (1967). Different perspectives on how to define implicit learning abound in the literature (Perruchet & Vinter, 2002; A.S. Reber, 1989; Shanks & St. John, 1994). Some of these perspectives are introduced further ahead in this thesis. As a general starting point, I will adopt the definition introduced above, according to which learning proceeds unintentionally (the participants do not actively try to learn the relevant regularities) and the knowledge is difficult to verbalize. Additional issues, such as whether the resulting knowledge is unconscious (different takes on this controversial issue are introduced further ahead in this thesis) and whether implicit learning reflects a separate learning system, are treated as empirical questions.

Implicit learning research extends into many different domains and is represented by research programs occupied with the question of how and what people learn in implicit learning situations, and how that learning affords sensitivity to the structure of different domains, (French & Cleeremans, 2002; Jimenez, 2003; A.S. Reber, 1989; Shanks, 2005; Stadler & Frensch, 1998). A number of more basic questions can be framed within this research field, questions that are of general relevance to psychology as a whole. For

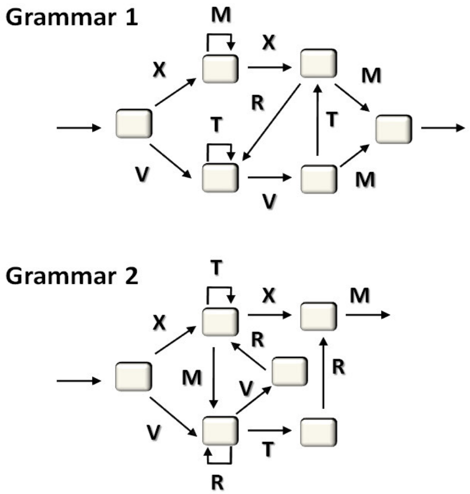

Figure 1. Two artificial finite-state grammars (adapted from Conway and Christiansen, 2006, and used in Study III in this thesis). Legal sequences are generated by entering the grammar from the left side and following the arrows until one reaches the exit on the right side. For example, in the top grammar XMMXRVM is legal (the consecutive Ms are due to the M loop, i.e., the arrow starting from and leading to the same state).

example, how do people generalize their knowledge to new instances that they have never encountered? Does unintentional learning give rise to a different kind of knowledge than active intentional learning (i.e., when participants are directly encouraged to extract regularities)? Can learning occur unconsciously, and what can be learnt unconsciously? How do people apply knowledge acquired through implicit learning? What is the functional role of consciousness?

In what follows I will shortly introduce some basic paradigms for studying implicit learning, particularly *artificial grammar learning* (AGL), the experimental paradigm of choice in all of the studies included in the current thesis (Study I, II, III, and IV). I will then introduce some of the questions and perspectives that have formed a basis for the included studies. Specifically, I will focus on two general issues, 1) the nature of the knowledge acquired in AGL, and 2) the conscious status of the knowledge acquired in AGL. This is then followed by a summary of the included studies and how they each illuminate some of the various issues raised regarding AGL and implicit learning generally. Finally, the specific studies follow as separate manuscripts.

Experimental Paradigms for Studying Implicit Learning

Artificial Grammar Learning

One of the most commonly used tasks in the study of implicit learning is AGL. In the typical version of this task the participants observe or memorize sequences of letters (or other symbols) that follow an underlying set of rules, without being told about the presence of rules. After this initial learning phase, the participants are informed that the sequences were generated by a set of complex rules and are shown *new* sequences, half of which follow the rules (grammatical) and half of which do not (ungrammatical). The task is to classify which sequences do or do not follow the rules. Usually, the participants can do so above chance levels (Pothos, 2007; A.S. Reber, 1967, 1989).

Most studies on AGL have used finite-state grammars that specify local sequential state-to-state dependencies (Figure 1). A number of studies have also investigated other kinds of grammars, such as biconditional grammars that specify non-local dependencies between symbols across a range of non-dependent intervening positions (e.g., Cock, 2005; Johnstone & Shanks, 2001; Kuhn & Dienes, 2005, 2006). For example, in a biconditional grammar, symbol X in position 1 may predict symbol Y in position 5 of a sequence, without having any predictive relation to symbols in other positions.

AGL constitutes an interesting paradigm for investigating implicit learning, because the stimulus domain is rather complex and dynamic. The sequences can be processed in many different ways (e.g., as wholes or as chunks) and there are many kinds of regularities available for learning, for example the frequency and position of chunks within sequences (Johnstone & Shanks, 1999; Knowlton & Squire, 1994; Pothos, 2007). The coordination of different kinds of learning and the selective application of knowledge in structured domains are important topics (e.g., Whittlesea, Brooks, & Westcott, 1994) and AGL is one of the many ways in which these can be brought under the looking glass.

Serial Reaction Time Task

The serial reaction time (SRT) task was introduced by Nissen and Bullemer (1987) and has been used extensively to investigate implicit learning (e.g., Destrebecqz & Cleeremans, 2001; Fu, Fu, & Dienes, 2008; Shanks & Perruchet, 2002; Wilkinson & Shanks, 2004). The task is to react as fast as possible to a stimulus on a screen. On each trial the stimulus appears on one of a number of possible locations and each location has a corresponding response button. The sequence of stimulus locations is not random, but follows an underlying sequence (either deterministically or with some specified probability). The participants´ response times decrease as training progresses and increase during blocks or trials when the underlying sequence is changed, showing that the participants become sensitive to the underlying structure of the training sequence.

Statistical Learning

Statistical learning (SL) is not usually framed as a direct method of studying implicit learning, but there are many similarities between, for example, SL and AGL (Conway & Christiansen, 2006; Perruchet & Pacton, 2006). SL was introduced in order to investigate the ability of infants to segment a continuous artificial speech stream into word-like units by computing transitional probabilities among sub-units in the speech stream (Saffran, Aslin, & Newport, 1996). The material used in SL is usually not specified directly by rules, but can be characterized by specific statistical relations among elements. Like AGL, SL proceeds without direct intention to learn the underlying regularities, which opens up the possibility for research comparing the two paradigms (Perruchet & Pacton, 2006).

Other Paradigms

In addition to the mentioned paradigms, there are many additional paradigms and methods that are about, related to, or relevant for the study of implicit learning. These include the study of *dynamic control systems* (e.g., Dienes & Fahey, 1995), *off-line learning* (e.g., sleep; Gomez, Bootzin, & Nadel, 2006), *invariant learning* (e.g., Kelly & Wilkin, 2006), *habit learning* (e.g., Bayley, Frascino, & Squire, 2005), *multiple-cue learning* (e.g., Lagnado, Newell, Kahan, & Shanks, 2006), *language regularities* (e.g., Perruchet & Peereman, 2004), *recognition memory* (e.g., Kinder & Shanks, 2001; Lotz & Kinder, 2006a), and *categorization* (e.g., Rouder & Ratcliff, 2006; Whittlesea & Leboe, 2000). Each of these areas may differ in various ways from what may be conceived as “prototypical” implicit learning situations (if there is such a thing), but they all reveal different aspects about sensitivity to the general structure of different domains. The interested reader may consult the cited sources above to gain a fuller understanding of these tasks and methods.

In what follows I will focus mainly on AGL and refer to other specific paradigms where this is of particular interest.

On the Nature of the Knowledge Acquired in AGL

What is the nature of the knowledge that affords sensitivity to general structure? More specifically, what kind of knowledge do participants acquire in AGL that allows them to distinguish between grammatical and ungrammatical sequences? Although posed as a relatively specific question, the answer has potentially wide implications because sensitivity to general structure is a basic feature of every-day life.

One answer to the question about the nature of the knowledge acquired in AGL is provided by the idea that participants acquire *abstract* knowledge of the rules of the experimental grammar, or of some set of rules approximating the grammar which is also consistent with and representative of the training items (A.S. Reber, 1967, 1989). In other words, the idea is that the participants become sensitive to structure by abstracting and representing that structure directly in a more centralized (rather than distributed) form, much like the classic idea that participants may learn a category by forming a prototypical or rule-based representation of the stimulus domain (Rosch & Mervin, 1975; Rouder & Ratcliff, 2006).

The idea of abstraction has been a subject of quite intense debate in AGL (Perruchet & Vinter, 2002; Shanks & St. John, 1994), much because of the additional notion, to be discussed further ahead in this thesis, that such abstraction is sometimes held to occur unconsciously and give rise to unconscious knowledge (A.S. Reber, 1989). In order to look closer at the abstraction theory it is necessary to specify the meaning of "abstraction" more carefully. In the next section I distinguish between three different senses of "abstract", each of which is a particular version of a more general sense of "abstract".

The Meaning of "Abstract"

In general, X is abstract with respect to Y when X is *drawn away* from Y in some sense (Litman & A.S. Reber, 2002). More specifically, Redington and Chater (1996) distinguished between three notions of abstract knowledge, each of which may be seen as a particular version of the more general notion above.

First, knowledge can be abstract in the uncontroversial sense of being separate from the actual physical stimulus. For example, an image can be encoded and stored as a face rather than as pixels of specific colors.

Second, knowledge can be abstract in the sense of being accumulated and *abstracted across exemplars*. For example, in AGL one line of theorizing holds that the processing of different sequences results in accumulation of knowledge stored as a set of rules representing the structure behind the sequences (A.S. Reber, 1989). Such knowledge is much more abstract in this second sense of abstraction than simply storing the sequences separately without further extraction of common features across sequences (Vokey & Brooks, 1992).

Third and finally, knowledge can be abstract in the sense of being divorced and independent from the specific surface properties of the training materials. Abstraction in this third sense may be said to lead to *surface-independent knowledge*.[1] Such knowledge can be successfully applied, not only in

1 Abstraction in the first and third senses may appear similar, but they are different. Abstraction in the first sense always occurs in the sense that incoming sensory information is always interpreted as something, for example, a sequence of letters. Abstraction in the third sense implies further abstraction into a format that, ultimately, enables knowledge application equally well for the original stimulus domain as for stimuli instantiated with different surface features.

the original training domain, but also in new domains where the same regularities are instantiated in different surface features (i.e., the symbols instantiating the grammar change, but the structure is preserved).

The second and third senses of abstraction specified above will form the basis for the following discussion of the nature of the knowledge acquired in AGL. It is important to note that these two senses of abstraction are conceptually independent from each other (cf. Redington & Chater, 2002). Knowledge may in principle be accumulated across exemplars and represented in a more centralized form than simply storing the exemplars separately (abstraction in the second sense above), while at the same time involving only the specific features that was encoded during learning (e.g., the letters "M", "V", "X", and so on; lack of abstraction in the third sense above). Conversely, knowledge may in principle be stored in terms of separate exemplars involving virtually no abstraction across exemplars (second sense of abstraction above), while at the same time involving abstraction into more general symbol features than those that were originally encoded (third sense of abstraction above).

Different theories of AGL are associated with different kinds and degrees of abstraction, both in the second and the third sense of abstraction mentioned above. For example, regarding the abstraction-across-exemplars issue some theories (e.g., A.S. Reber, 1989) hold that participants unintentionally abstract rules across the exemplars during learning, while other theories (e.g., Vokey & Brooks, 1992) hold that encoded sequences are stored during learning with virtually no unintentional abstraction across exemplars during the learning phase. In what follows these different theories are introduced and briefly discussed, starting with the debate concerning the second sense of abstraction (across exemplars) and then proceeding to the third sense of abstraction (surface-independence).

On the Nature of Abstraction Across Exemplars in AGL

Theories of AGL differ in the extent to which they associate learning with the computation of abstract representations during learning in the sense of abstraction across exemplars. In what follows I focus on how five general accounts of AGL relate to the abstraction-across-exemplars issue, namely *rule abstraction* accounts, *similarity* accounts, *fragment* accounts, *statistical* accounts, and the *episodic-processing* account. (The fragment and statistical accounts are treated within the same section because they both emphasize sensitivity to chunks (parts) of sequences.) This is then followed by a section on the flexible and heuristic nature of classification in AGL, an issue which is crucial in order to reach an understanding of what participants learn and express in AGL and implicit learning situations generally.

Rule Abstraction Accounts

According to A.S. Reber (1967, 1989; see also Chang & Knowlton, 2004; Knowlton & Squire, 1994, 1996; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Marcus, Fernandes, & Johnson, 2007) sensitivity to structure in AGL results from (unconscious) abstraction of regularities across exemplars in the form of rule knowledge. The application of this pre-computed rule knowledge forms the basis of grammaticality judgments concerning new sequences at test. By *pre-computed*, I simply wish to emphasize that this account assumes that the knowledge that affords sensitivity to general structure is abstracted during training into the form in which it is later used.

Although abstract rules can account for participants´ ability to distinguish between grammatical and ungrammatical sequences, there are a number of challenges facing the idea that the participants represent general structure directly in the form of abstract rules. For example, A.S. Reber´s (1989) rule account is framed in terms of high-threshold theory, so that the participants either know the grammatical status of a sequence or they guess. However, analyses of receiver-operating characteristics in AGL show no evidence of a high-threshold process, but instead point to results consistent with a *continuous* underlying memory variable as conceptualized within signal detection theory (Kinder & Assmann, 2000; Lotz & Kinder, 2006a, 2006b). Furthermore, as we shall see, many studies of AGL reveal that the knowledge acquired in incidental learning situations can be applied flexibly and heuristically in ways that do not fit naturally with the idea that the knowledge was automatically abstracted and directly represented during training (Johnstone & Shanks, 2001; Kinder, Shanks, Cock, & Tunney, 2003; Lotz & Kinder, 2006a; Vokey & Higham, 2005; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998). In addition, as discussed further ahead, although the knowledge acquired in AGL affords sensitivity to general structure, it does not seem to be the case that the knowledge is constituted by representation of structure *per se* (Higham, 1997a; Whittlesea & Dorken, 1993).

Before going into these studies in more detail, it is informative to consider some other alternative AGL accounts that rely to a lesser extent on pre-computed abstract knowledge, but instead capitalize on *similarity* of different kinds.

Similarity Accounts

Brooks (1978) suggested that instead of being a result of abstraction of general structure across exemplars during training, sensitivity to general structure may emerge as a by-product of the storage of *specific exemplars*. That is, during training the participants are assumed to store specific exemplars, for example, entire letter sequences. At test, new sequences are judged as grammatical or ungrammatical on the basis of their similarity to training sequences, not on the basis of abstract rules. The basic idea of sensitivity to general structure as an emergent property of rather specific representations is central to a wide variety of models in both the memory and categorization literature (e.g., Hintzman, 1986; Nosofsky, 1986, Nosofsky & Zaki, 1998).

Without controlling for the similarity between grammatical and ungrammatical test sequences in AGL, grammatical test sequences tend to be more similar to the training sequences than ungrammatical test sequences are. In effect, responding on the basis of similarity between a test sequence and the training sequences provides an indirect route to sensitivity to general structure *without* any direct computation of abstract knowledge across exemplars during the learning phase. The statistical properties of the training materials are indirectly preserved through the distributed storage of specific exemplars.

Vokey and Brooks (1992) operationalized similarity as *item-specific similarity* (sometimes also referred to as "edit distance"), defined as the number of transformations required to change a test sequence into the most similar specific training sequence. Orthogonal to grammatical and ungrammatical test sequences, *close* test sequences differed from the most similar training sequence by one letter, and *far* test sequences differed from the most similar training sequence by at least two letters. Vokey and Brooks found additive effects of both grammaticality and item-specific similarity, the latter indicating that close test sequences were claimed to be grammatical by the participants more often than far test sequences. Although it could be argued that the residual effect of grammaticality reflected abstract rules, Vokey and Brooks favoured the interpretation that the remaining grammaticality effect was a result of matching between a test sequence and a "chorus of instances" (i.e., several training sequences), while the item-specific similarity effect resulted from matching between a test sequence and a specific training sequence (cf. Hintzman, 1986).

The notion of item-specific similarity has been criticised on the basis that it may be reducible to fragmentary knowledge (see the next section on *Fragment and Statistical Accounts*; Perruchet, 1994; Knowlton & Squire, 1994). However, effects of item-specific similarity have been observed under a variety of conditions in which an explanation in terms of fragmentary knowledge is less likely (Higham, 1997a, 1997b; Lotz & Kinder, 2006a, 2006b; Pothos & Bailey, 2000; Vokey & Higham, 2005).

Item-specific similarity refers to a match between a specific test item and a specific training item. A different version of exemplar knowledge is implemented by Robert Nosofsky´s (1986) *generalized context model* (GCM). In this model, classification is based on the psychological similarity in a multidimensional space between a test item and *all* of the training items. The similarity space is based on empirical similarity ratings between items and not on a pre-defined assumed similarity metric, which adds plausibility to the model. The model has been very successful in studies on category learning and was tested in several AGL experiments by Pothos and Bailey (2000). The results showed that similarity as embodied in the GCM is a significant predictor of grammaticality judgments in AGL, over and above that of several other predictors (e.g., item-specific similarity and fragmentary knowledge). Further investigation of exemplar models in the context of AGL would constitute an interesting research topic (for related research on structured dot patterns, see Zaki and Nosofsky, 2007).

Fragment and Statistical Accounts

An additional source of sensitivity to general structure in AGL is *fragmentary knowledge* (or *chunk knowledge*). According to this account, participants become sensitive to the distribution of fragments or chunks within the set of training sequences (e.g., bigrams and trigrams, which refers to parts of sequences of two and three symbols). At test, participants may classify sequences as grammatical depending on the extent to which the distribution of chunks in a test sequence overlap with the distribution of chunks in the training set (Johnstone & Shanks, 1999, 2001; Knowlton & Squire, 1994; Perruchet & Pacteau, 1990; Servan-Schreiber & Anderson, 1990).

In support of a fragment account, Perruchet and Pacteau (1990) found that subjects trained on bigrams classified entire test sequences with levels of accuracy similar to subjects trained on entire sequences. Perruchet, Vinter, Pacteau, and Gallego (2002) used a segmentation task in the context of AGL (marking the natural segmentation points of sequences) and found that the number of formed chunks did not differ before and after training, but the number of *different* chunks were lower after training than before. This is consistent with fragmentary models in which re-occurring chunks are strengthened (although models which assume formation of increasingly larger chunks during training are at odds with the results, such as Servan-Schreiber and Anderson, 1990).

One similarity metric often used in the assessment of fragmentary knowledge is *chunk strength*, which is defined as the average frequency with which the bigrams and trigrams of a test sequence occurred in the entire training set as a whole, regardless of position within a sequence (Knowlton & Squire, 1994). Many studies have found effects of both grammaticality and chunk strength, which has led to the suggestion that AGL gives rise to both abstract rule knowledge and distributed chunk-based knowledge (Chang & Knowlton, 2004, Knowlton & Squire, 1994; Meulemans & Van der Linden, 1997). However, using within-subjects regression analyses (Lorch & Myers, 1990) Johnstone and Shanks (1999) found that grammaticality was not a significant predictor of classification judgments when a variety of predictors were taken into account (e.g., whether a test sequence contained a novel chunk or not and whether it contained a chunk in a novel position or not), suggesting that effects of grammaticality may be a result of confounds between grammaticality and other kinds of predictors (cf. Kinder and Assmann, 2000).

Different fragmentary accounts based on the formation of chunk knowledge differ in various ways, but they usually assume that the strength of stored chunks may change over the course of learning as a by-product of processing of different exemplars (Knowlton & Squire, 1994; Perruchet & Vinter, 2002; Servan-Schreiber & Anderson, 1990).[2] Thus, the knowledge may be said to be partly abstract in the sense that it accumulates and is strengthened over the course of encoding chunks in a variety of different entire sequences. Chunks stored in memory may plausibly be viewed as less abstract than rules accumulated during the course of learning, but more abstract than the storage of specific exemplars (in the latter case general structure exhibited at test emerges *only* and *wholly* at test). Note that the issue here is not the computational sophistication of the learning mechanism. Chunking may proceed through quite simple associative learning mechanisms (Perruchet & Vinter, 2002) and sensitivity to the distributional statistics of a set of regularities may also emerge as a result of degenerate limited encoding (Whittlesea & Dorken, 1993).

As discussed by Boucher and Dienes (2003; cf. Perruchet, 2005; Perruchet & Pacton, 2006; Perruchet & Peereman, 2004) sensitivity to the distributional features of the training materials in AGL can occur both through direct encoding and storage of chunks (Perruchet & Vinter, 2002; Servan-Schreiber & Anderson, 1990) and through *statistical computations* as embodied in connectionist models, most notably the simple recurrent network (SRN) model (Elman, 1990; Kinder, 2000; Kinder & Shanks, 2001; Boucher & Dienes, 2003). The SRN contains an input layer where a stimulus is presented, a hidden layer where an internal representation is formed, and an output layer where a response is generated. In addition, it contains a context layer which is connected to the hidden layer. The units of the context layer represent the hidden layer activation at the previous time step (t -1). At each time t, the network is presented with one symbol of a sequence and then tries to predict the next letter. The weights in the network are adjusted through backpropagation after each prediction attempt. The context layer allows for the network to predict the next symbol in a sequence on the basis of, not just the immediate previous symbol, but several previous symbols back in time. In effect, as training progresses the network predicts the next symbol on the basis of increasingly higher-order sequential statistical dependencies (which may be indirectly reflected in chunk strength measures). Prediction accuracy with respect to test sequences can be transformed into a grammaticality decision by way of a decision rule (Dienes, 1992; Kinder & Shanks, 2001).

The SRN has been shown to account for a variety of results in AGL and may be considered as one of the most successful computa-

2 As noted by Pothos (2007), in competitive chunking (Servan-Schreiber & Anderson, 1990) chunks representative of the structure of a domain are formed in terms of higher-order chunks on the basis of co-occurrence of lower-order chunks or individual elements. In PARSER (Perruchet & Vinter, 2002) on the other hand, chunks are formed on the basis of both co-occurrence and interference (the strength of AB is reduced if A or B occur elsewhere as well, e.g., in AC or CB).

tional models of AGL (Dienes, Altmann, & Gao, 1999; Kinder, 2000; Kinder & Shanks, 2001). However, Boucher and Dienes (2003) found that the competitive chunking model (Servan-Schreiber & Anderson, 1990) reflected participants´ susceptibility to interference more characteristically than the SRN. Kuhn and Dienes (2008) found that a buffer network with simultaneous coding of many previous time steps reflected participants´ sensitivity to non-local dependencies better than the SRN. Also, although not directly related to AGL, Perruchet and Peereman (2004) found that a chunking model, PARSER, accounted for participants´ word-likeness judgments on the basis of statistical relations between phonemes better than did the SRN (the SRN was mainly sensitive to transitional probabilities, while PARSER was sensitive to more sophisticated association measures).

Although chunking models and the SRN imply different kinds of learning, the resulting knowledge may in both cases be viewed as less abstract (and the learning process itself less abstractionist) than rule learning. In neither of the two kinds of models above is there any direct centralized representation of the general structure of a domain. Rather, the structure of a domain is distributed across and reflected in the stored chunks and their strengths (chunk models) or in the weights of a network (simple recurrent network).

As reviewed so far, a number of accounts exist that rely on less abstract knowledge, in the sense of abstraction over exemplars, than that implied by direct abstraction of the rules governing the structure of a domain (for a more detailed review, see Pothos, 2007). However, there is an additional account according to which the nature of the knowledge acquired and exhibited in AGL is very much context-dependent. This account is usually referred to as the episodic-processing account.

The Episodic-Processing Account

According to the *episodic-processing* account of implicit learning (Higham, Vokey, & Pritchard, 2000; Jamieson & Mewhort, 2005; Neal & Hesketh, 1997; Shanks, Johnstone, & Kinder, 2002; Whittlesea & Dorken, 1993; 1997; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998; for a more general view, see Bruce Whittlesea´s, 2003, SCAPE account) the products of learning consist in episodic representations that incorporate processing experiences, not just representations that preserve objective structural relations. The content of the stored episodes is not simply a function of intention to learn, or of the nominal status of a task, or of the specific regularities embodied within a stimulus set. Instead, the nature of the episodic representations is a *combined* function of the participants´ intentions, processing strategies, expectations, learning history, and the stimulus structure of the domain (Whittlesea & Wright, 1997). In other words, there is no general fact of the matter concerning what participants learn in AGL and other implicit learning situations. Instead, it all depends. For example, if the learning task directly or indirectly encourages abstract processing, then abstract learning may very well emerge, not because implicit learning is dependent upon a special implicit learning system, but because the task as a whole encourages a particular storage and use of episodic representations.

In claiming that learning depends on many situational factors, the episodic-processing account is incompatible with the claim that automatic unconscious abstraction is the reason for sensitivity to structure. Chronic abstraction is not at the core of the episodic-processing account. Rather, implicit learning of structure is more a result of using episodic knowledge in unanticipated ways. For example, as noted by Vokey and Brooks (1992), sensitivity to general structure can be explained through pooling of several exemplars (or episodes) even though any representation of general structure is not directly stored.

The episodic-processing account may be viewed as a variant of the similarity accounts reviewed in the *Similarity Accounts* section. However, the episodic-processing account is less constrained and more flexible than the similarity accounts, as the nature of the episodic representations formed during learning and how these representations are later utilized can vary to a large extent depending on context in the former account. The most important points to note regarding the episodic-processing account are that it emphasizes 1) preservation of episodic knowledge (rather than automatic abstraction of structure), 2) flexibility of learning and application of knowledge (rather than specific kinds of knowledge being learned or triggered in specific tasks), and 3) a conceptualization of implicit knowledge as an indirect consequence of processing (not pre-computed structural knowledge). A number of studies serve to illustrate these points.

Whittlesea and Dorken (1993, Experiment 1) trained participants on two grammars, requiring the participants to spell sequences from one grammar and pronounce sequences from the other grammar. At test, each participant pronounced half of the sequences from each of the grammars and spelled the rest. The results showed that classification was under control of the correspondence between the match of processing operations conducted during training and test. For example, participants were less likely to endorse sequences from the spell grammar if the new sequences shown at test from that grammar were pronounced. These results constitute evidence that the knowledge acquired in AGL is not purely structural, but consist in episodic representations that preserve processing experiences. Studies by Wright and Whittlesea (1998) point to similar results in invariant learning, where the hidden rule consists in a specific invariant feature in the training exemplars (the invariant feature might be an odd-even pattern in a sequence of numbers, rather than the more complex rules in AGL). Furthermore, Higham (1997a) found that making letter sequences pronounceable in AGL had an impact on classification as compared to less pronounceable sequences, even though the structural information was equivalent in the two conditions. This shows that AGL is not unselective and not purely structural.

Kinder et al. (2003) showed that classification in AGL is flexible and can proceed on the basis of different heuristics, depending on what kind of processing strategies are adopted at test, indicating flexible application of knowledge in AGL. The authors used both recognition and classification tests and showed that although the participants were biased toward using different heuristics depending on instructions to classify on the basis of grammaticality or recognize whether sequences had been shown before during the experiment, it is possible to manipulate which heuristics participants use in both classification and recognition without changing the nominal status of the test task (this study is discussed further in the *Heuristics in AGL* section). In other words, application of knowledge is not specifically tied to the nominal status of a task. In addition, the results support processing accounts of dissociations between recognition and classification (according to which such dissociations arise because of specific task demands) and cast doubt on the necessity of postulating different memory systems (e.g., implicit vs. explicit memory systems) on the basis of such dissociations (cf. Dunn, 2003; Kinder & Shanks, 2001, 2003; Love & Gureckis, 2007; Shanks, 2005; Shanks & Perruchet, 2002; Tunney & Shanks, 2003b; Whittlesea & Price, 2001).

Johnstone and Shanks (2001) found that participants could not learn a biconditional grammar under standard implicit learning instructions. Instead, the participants classified on the basis of chunk strength. However, when given a training task that encouraged discovery of the relevant features, the participants were able to learn the grammar. In other words, the nature of the specific training task, not just the structural features of the learning environment, dictates what participants learn in AGL.

Pothos (2005, Experiment 2) found that in a modified AGL task, using names of cities as symbols, classification accuracy was lower when participants´ expectations about stimulus structure was in conflict with the actual structure compared to when expectations were congruent with the actual structure of the stimuli. The participants were told a cover story saying that the sequences to be shown were the routes of a travelling salesman, and that the routes were planned so as to be maximally efficient with many visits between nearby cities. In the congruent condition the structure of the stimuli was in line with the cover story, but in the incongruent condition it was not. The results suggest that AGL is susceptible to explicit expectations and not purely structural and unselective.

Jamieson and Mewhort (2005) investigated AGL in the context of a short-term memory task rather than classification. On each trial the participants viewed a sequence and then attempted to recall it (cf. A.S. Reber, 1967, 1969). The authors found that the number of recall errors decreased over trials to a larger extent when the sequences being memorized were from a more constrained grammar as compared to a less constrained control grammar. The authors went on to show that this memorization benefit was most likely *not* due to learning of the grammar. Rather, the benefit for more constrained grammatical sequences was due to information reduction in terms of organizational redundancy, which is a measure of the degree to which the different units of a sequence, encoded in a particular way, predict each other. For example, the measure captures why RGBYRGBP is easier to memorize than RGBYPRGB when encoded as chunks of four sequential symbols each, since such chunking increases the redundancy of the entire sequence differently in the two cases. Jamieson and Mewhort (2005) found that memorization advantages for particular sequences were well described by the organizational redundancy as given by the particular encoding operations used. Encoding operations that gave rise to lower organizational redundancy were more difficult to memorize. Thus, in line with the episodic-processing account, rather than learning an abstract grammar the participants capitalized on the redundancy of particular sequences as given by the specific ways in which the sequences were encoded.

As reviewed so far, a number of studies strongly indicate that AGL is selective, sensitive to specific encoding operations, and context-dependent. In addition, the episodic-processing account emphasizes the heuristic nature of classification (Lotz & Kinder, 2006a; Kinder et al., 2003; Whittlesea & Leboe, 2000). The idea is that the representations laid down during training may be used flexibly and give rise to dissociable patterns of performance, even though everything is grounded in one and the same representational system. Taken together, these heuristics offer a potentially unitary account of AGL by showing that different heuristics are used under different circumstances.

Heuristics in AGL

Most AGL studies have focused on conditions during training, treating the test phase as a more or less neutral index of what knowledge participants have gained. However, it is quite clear that conditions during the test phase influence what and how knowledge is applied (Helman & Berry, 2003; Whittlesea et al., 1994). This fits well with the episodic-processing account, which views application of knowledge as heuristic and implicit knowledge as partly constructed in the moment it is applied (Whittlesea, 2003).

Whittlesea and Leboe (2000) suggested that participants may rely on at least three different heuristics in classification, namely the *resemblance* heuristic, the *generation* heuristic, and the *fluency* heuristic. These heuristics are assumed to operate on the content and effects of an episodic memory system which preserves processing experiences.

The resemblance heuristic involves relying on general information in classification, such as *non-specific* similarity derived from the distribution of chunks in the entire set of training sequences (for example, the test sequence MVXXL might be made up of chunks that were encountered frequently during training). In contrast, the generation heuristic involves relying on specific information, such as specific training sequences (for example, the test sequence MVXXL might trigger a specific very similar training sequence, say MTXXL) (Lotz & Kinder, 2006a; Vokey & Brooks, 1992). Finally, the fluency heuristic involves relying on the surprising ease of processing a test sequence as a cue when classifying items (Kinder et al., 2003, Newell & Bright, 2001; Whittlesea & Dorken, 1993). Since grammatical test items are processed more fluently than ungrammatical test items after exposure to grammatical training items (Buchner, 1994) relying on a fluency heuristic may be considered an adaptive strategy (although, as with all heuristics, it can also lead to systematic errors).

It is plausible to suppose that fluency can be derived from both specific and non-specific similarity (Lotz & Kinder, 2006a). However, whether participants rely on that fluency or whether they instead rely on the familiarity provided by resemblance or on the contextual detail provided by generation of information may depend on a variety of factors.

In order to demonstrate the use of a fluency heuristic in AGL, Kinder et al. (2003) used a perceptual clarification procedure in the test phase, so that some sequences clarified faster than others on screen. The participants endorsed fast-clarifying test sequences as grammatical more often than slow-clarifying test sequences, providing evidence for use of the fluency heuristic. However, when asked to ignore the fluency manipulation, the participants were unaffected by fluency. Also, when asked to recognize which sequences had been shown during the training phase, the participants were unaffected by fluency when the training sequences were shown at test, but not when the training sequences were excluded from the test phase.

Kinder et al. (2003) concluded that use of the fluency heuristic is flexible and adaptive, and may change as a function of processing strategies during the test phase.[3] When the training sequences were included at test, the participants who received recognition instructions relied on *analytic processing* (and processed details of the sequences in order to generate contextual detail), which prevented experiences of fluency. However, when the training strings were removed from the test phase, the participants switched to *non-analytic processing* (and processed the sequen-

[3] The adoption of particular processing strategies (or combinations thereof), such as analytic and non-analytic processing, does not necessarily have to be a result of an active conscious choice on behalf of the participants. Thus, the word “strategy” may be somewhat misleading given that its every-day application often includes that a strategy is actively selected. In the present context, particular ways of processing information can be partly triggered by internal and external cues, the impact of which the participants do not necessarily reflect actively upon to any larger extent. The adoption of processing strategies may perhaps be viewed as resulting from a set of combined factors, some of which the participants are aware of and some of which they are less aware of.

ces as wholes), enabling experiences of fluency (cf. Whitttlesea and Price, 2001, for a similar perspective in a different domain). Study I in this thesis further investigated the role of fluency and its relation to non-analytic processing in AGL by using a masked priming procedure to manipulate fluency (e.g., Jacoby & Whitehouse, 1989) and speeded responding to encourage non-analytic processing.

Lotz and Kinder (2006a) compared receiver-operating characteristics (ROCs) associated with classification and recognition in AGL. When the training sequences were included at test, the ROCs differed between classification and recognition, consistent with the results obtained by Kinder et al. (2003). By investigating the distribution of different underlying similarity sources (specific and non-specific similarity), the authors concluded that their results were consistent with a greater reliance on the generation heuristic (remembering specific sequences) in the recognition group and a greater reliance on resemblance or fluency in the classification group.

The resemblance heuristic may appear incompatible with the storage of episodic representations, since the resemblance heuristic is associated with reliance on general, rather then specific, information. However, the whole question turns on in what sense and when general information is abstracted or constructed out of more specific or episodic information. According to the episodic-processing account, the content of episodic representations can be abstract in each of the senses described in the current thesis if the task at hand encourages a particular kind of processing relevant for a particular type of abstraction. Furthermore, abstraction over a set of episodic representations (similar to the "chorus of instances" referred to by Vokey and Brooks, 1992) may occur at *test*, when the direct need and motive for abstraction becomes apparent in most implicit learning experiments. Thus, reliance on general information does not in itself imply that the general character of the information was directly and automatically stored during the learning phase.

Wright and Whittlesea (1998, Experiment 5) used an invariant learning paradigm in order to demonstrate that sensitivity to general information may be a result of performing abstraction at test. The authors constructed training materials with a highly abstract property that was not mentioned to the participants, namely that none of the exemplars (four-digit numbers, e.g., 6945) presented a regular pattern (e.g., 1234). After exposure to these exemplars, the participants received a set of various test pairs. Each test pair consisted of a similar and a dissimilar (relative to specific training numbers) four-digit number (see Wright and Whittlesea, 1998, p. 414, for details regarding the construction of these materials). Furthermore, half of the similar numbers contained a salient regular pattern (e.g., 1234) while half did not. The task was to say which of the numbers had been seen earlier in the exposure phase (when in fact none had been seen in the exposure phase).

The results showed that for test pairs that contained similar irregular and dissimilar irregular numbers participants did not distinguish between the two types of numbers in their recognition judgments. However, when the test pairs contained similar regular and dissimilar irregular numbers participants claimed the dissimilar numbers to be old more often than they claimed the similar regular ones to be old. Hence, the participants were sensitive to the abstract property of lack of regularity in the training materials. It can not be *proven* that this property was not abstracted during the incidental exposure phase, but, as the authors point out, "if one wants to imagine that subjects directly worked out, during the training phase, that all stimuli were irregular, then one also has to imagine that they also learned all manner of other properties, which were also invariant during the training phase" (p. 414). Clearly, ubiquitous abstraction of that kind does not seem plausible.

Instead of being pre-computed, the knowledge of the abstract property in the training materials may have existed as a *potential* after training, and was created on the fly during testing. Participants then probably became aware that some of the test items were regular (e.g., "Hmm, 1234…that looks different from the others, ah, it is regular.") and that they should have been able to recall such regular items if they had seen any of them before (e.g., "How could I have seen 1234, which is salient, and not remember it?"). There are many other abstract dimensions that the participants could have been become sensitive to during testing had the test materials been different. This knowledge may be described as "implicit", "latent", or as a "potential", but it is not pre-computed.

On the Nature of Abstraction Across Exemplars in AGL: Summary

A variety of theories have been offered in order to explain how people display sensitivity to general structure in AGL. Abstract rule theories hold that participants abstract the general structure of a domain across different exemplars in the form of rules during the learning phase. Similarity theories claim that classification of new test items is based on some form of similarity to training exemplars. Theories that emphasize fragment knowledge (e.g., chunking models and the simple recurrent network) involve a kind of abstraction across exemplars, but nevertheless not the kind of direct abstraction involved in rule theories.

Support can be found for both exemplar and fragment knowledge in AGL, although the latter is less controversial than the former. There is little, if any, direct evidence for unintentional abstraction of rules in AGL. Instead, in line with the episodic-processing framework, abstract rule knowledge can be formed but this does not seem to occur in a way that is independent from task constraints, task demands, and intentions.

The coordination of different kinds of knowledge in AGL (both in terms of acquisition and application) may be understood through the episodic-processing framework in combination with different kinds of heuristics. According to this view, learning in AGL does not reflect an unselective stimulus-driven separate implicit learning system. Rather, learning and classification in AGL and other implicit learning situations is viewed as a combined function of many different factors, for example intentions, expectations, stimulus domain, and processing strategies. Furthermore, knowledge is not automatically abstracted across exemplars into a more centralized form. Instead, when and whether such pooling of episodic representations occurs depends on task demands.

On the Nature of Abstraction Across Surface Form in AGL

In addition for knowledge to be abstract in the across-exemplars sense, knowledge can also be abstract in the sense of being divorced and independent from the specific instantiation used during training. Abstract knowledge in this latter sense is often referred to as *surface-independent knowledge* (Redington & Chater, 1996, 2002). In contrast, when knowledge is represented only in the form of surface features encountered during training it is said to be *surface-dependent*. As reviewed in what follows, a great deal of research has been devoted to investigating to what degree the knowledge acquired in AGL is surface-independent. Claims for surface-independence are controversial partly because they sometimes imply a sophisticated cognitive unconscious that performs the required abstraction automatically. In effect, it is mainly *automatic* abstraction toward surface-independence during training that is controversial.

Surface-Independent Knowledge

The main reason for invoking surface-independent knowledge in the discussion of AGL is that participants can *transfer* their knowledge between different domains in AGL. For example, even if the participants are trained on tone sequences, they can classify new sequences above chance levels even if the test sequences are instantiated as letter sequences instead of tones (Altmann, Dienes, & Goode, 1995; Manza & A.S. Reber, 1997). That is, the knowledge acquired in AGL does not only seem to afford generalization to new sequences within the same domain, but also to new sequences in new domains involving new stimulus forms. In effect, the knowledge acquired in AGL has been argued to be abstract in the sense of being surface-independent (Chang & Knowlton, 2004; Knowlton & Squire, 1996; Manza & A.S. Reber, 1997; A.S. Reber, 1969, 1989), for example in the form of algebraic rules operating over variables (Marcus, 2001; Marcus et al., 1999).

Transfer of knowledge to new stimulus domains in AGL is a robust phenomenon (Redington & Chater, 1996, 2002). In effect, the knowledge acquired in AGL has to be formed in a way that allows for application in new stimulus domains. However, a phenomenon equally robust as above-chance transfer performance is the so called *transfer decrement* effect, which refers to the fact that transfer performance is generally lower than same-domain performance (for discussion, see Perruchet & Vinter, 2002). The observation of transfer decrement has lead to a variety of accounts of the knowledge and mechanisms underlying transfer in AGL that do not presuppose knowledge automatically abstracted into a surface-independent format during training. These accounts are important because they demonstrate that the main evidence for automatic abstraction toward surface-independence in AGL, namely transfer to new domains, can be accounted for without assuming that knowledge is automatically abstracted into an abstract surface-independent form. Some of these accounts are reviewed in the next section.

Accounting for Transfer without Automatic Abstraction during Training

Brooks and Vokey (1991) suggested that transfer across domains can occur through the use of *abstract analogies*, involving comparison of test sequences to specific training sequences on an abstract level. For example, the sequence MVXXRV is identical to the sequence QTSSPT with respect to the repetition patterns embodied across the sequences. Both sequences contain a first element, followed by a different element, followed by two identical elements different from the first two elements, and so on. Brooks and Vokey (1991) found effects of abstract item-specific similarity in a transfer version of AGL, suggesting that the participants may have partly relied on abstract analogies. Although such analogies are referred to as abstract, it is important to point out that abstraction across surface forms may occur *at test*, when the direct need for abstraction arises. The knowledge laid down during training does not have to be surface-independent.

Vokey and Higham (2005) found transfer even with a randomly changing transfer paradigm, where a new mapping of symbols in the grammar is used for each test item. Since repetition information is partly preserved in such a paradigm it is possible to use abstract analogies. The authors also found (somewhat ironically for the abstract rules account) that compared to same-domain performance, the grammaticality effect (endorsing grammatical over ungrammatical sequences) was reduced but the item-specific similarity effect was unaffected by randomly changing transfer. Gomez, Gerken, and Schvaneveldt (2000) only found evidence for transfer when the grammar that participants were trained on contained repetitions of symbols, suggesting that symbol repetitions are crucial for transfer performance. These results are consistent with the use of abstract analogies as a basis for transfer performance in AGL.

Lotz and Kinder (2006b) conducted within-subjects regression analyses including various similarity predictors in a transfer version of AGL and found that repetition patterns was a significant predictor of classification. Specifically, the authors investigated both global and local repetition patterns. The former refers to repetition of symbols in an entire sequence and corresponds to the kind of repetitions discussed above in the context of abstract analogies, whereas the latter only refers to repetitions of adjacent symbols. For example, the local repetition pattern of the sequence MVXXRV is 00100, where 0 indicates that a symbol is different from its immediate predecessor and 1 indicates that it is identical. In contrast, the global repetition pattern of the sequence is coded as 012201, where 0 indicates a non-repeating symbol (M and R in the example sequence) and 1, 2, and so on, indicate specific repeating letters across the sequence (V and X in the example sequence). Note that two different symbols may both be indicated by 0, if none of them are repeated in the sequence. Lotz and Kinder (2006b) found evidence for reliance on both global and local repetition patterns in transfer AGL (i.e., when the surface form is changed between training and test). Other types of similarity predictors (e.g., chunk strength) had no influence in transfer AGL, but only in non-transfer versions of the task. Both local and global repetition patterns imply some form of abstraction toward surface-independence, but the important question is when and how this abstraction takes place. Repetition patterns (in terms of elementary perceptual or motoric “same-different” relationships) may be stored directly during encoding and would thus not involve any gradual cognitive abstraction during learning (Lotz & Kinder, 2006b; McClelland & Plaut, 1999; Perruchet & Vinter, 2002). Also, as already mentioned and expanded further in what follows, abstraction may occur intentionally at test rather than automatically during learning.

Redington and Chater (1996) evaluated various “toy models” of transfer performance based on surface-dependent fragment knowledge. The models were referred to as toy models because they were not intended as psychologically plausible models, but simply as demonstrations of the fact that transfer to new domains can occur on the basis of surface-dependent knowledge. As in the account of Brooks and Vokey (1991), Redington and Chater assumed that abstraction towards surface-independence occurs at test rather than automatically during training. This is reflected in the models, which attempt to induce possible mappings between surface statistics in the training materials and test sequences in the new domain. For example, say that a number of test sequences have been memorized in the training domain (letter set 1) and a new test sequence is encountered in the new domain (letter set 2). Then, the new test sequence may be rejected because it is not possible to find a mapping between memorized bigrams in the training domain and the bigrams in a new test sequence in the new domain (for details, see Redington and Chater, 1996). The important point to bear in mind is that the simulations in Redington and Chater (1996) show that transfer across surface forms can occur at test on the basis of surface-dependent representations formed during training.

Dienes et al. (1999) presented a connectionist model of AGL (an extended version of the simple recurrent network) that was able to fit the classification judgments of participants in transfer domains. The model works by indirectly mapping statistical dependencies between domains and is based on statistical learning of regularities. It differs from abstract analogies in that it is predicated on learning of statistical dependencies and assigns no special role to symbol repetitions within sequences. Again, the important point to bear in mind is that surface-dependent representations are formed in the model during training. The processes responsible for transfer across domains are activated at test. Tunney and Altmann (2001) found evidence both for transfer based on statistical learning and transfer based on abstract analogies. The authors found that knowledge of simple first-order statistical dependencies (the relation between two adjacent symbols) could be transferred to a new domain without invoking abstract analogies. Furthermore, the two modes of transfer could be dissociated in that transfer based on statistical knowledge was sensitive to changes in the distributional statistical properties of the training domain while transfer based on abstract analogies was relatively unsensitive to such distributional changes. However, the grammars used by Tunney and Altmann (2001) were relatively simple and it is not clear to what extent transfer can be based on statistical knowledge when using more complex grammars. As already mentioned, Gomez et al. (2000) found no transfer with more complex grammars containing no symbol repetitions within sequences.

The accounts reviewed so far in this section all illuminate ways in which transfer can be achieved on the basis of surface-dependent knowledge without automatic abstraction toward surface-independence during incidental learning of regularities. In effect, above-chance transfer performance does not by itself constitute evidence in favour of automatic abstraction during the learning phase in AGL. Instead, as discussed next, the degree of surface-independence associated with the representations formed in AGL may be a function of task demands, pointing to the relevance of an episodic-processing account (discussed previously in connection with abstraction across exemplars) in transfer AGL.

Episodic Processing and Transfer

The accounts in the previous section suggest ways in which transfer can occur in AGL on the basis of surface-dependent representations. In addition to these accounts, the more general episodic-processing account holds that the formation of surface-independent representations is a function of context, intentions, and task instructions, rather than a function of the automaticity associated with an implicit learning system triggered by the presence of complex regularities. In effect, surface-independent representations may very well form if the training or test task (directly or indirectly) encourages abstract processing. Thus, the episodic-processing account predicts that transfer performance in AGL should vary with the kind of processing encouraged by the task in the current context.

Whittlesea and Dorken (1993, Experiment 5) found that transfer to a new domain in AGL was better when the training task encouraged processing of repetitions within sequences. As already discussed, repetition information is central to transfer performance in AGL. In effect, the results of Whittlesea and Dorken (1993) indicate that the formation of knowledge relevant for transfer in AGL is dependent on the processing strategies triggered by the task, rather than the knowledge being formed automatically through a specialized implicit learning system.

A similar point can be made with respect to the test phase in AGL. If the formation of abstract knowledge does not occur automatically, but instead occurs as a function of intention and motives to engage in processes relevant for such abstraction, then transfer performance should be better when the test phase instructions point directly to the need for some form of abstraction compared to when the instructions do not. In line with this prediction, Newell and Bright (2001) found above-chance transfer performance in AGL when the test task was to classify new sequences, where the presence of underlying regularities was mentioned between training and test. In contrast, there was no effect of grammaticality (the grammatical status of test sequences) when an indirect test task (preference judgments) was used that did not mention the presence of underlying regularities. Preference judgments were only affected by grammaticality when the same domain was used at both training and test, in which case there was no need for surface-independent representations.

In line with an episodic-processing account, the studies discussed in this section indicate that explicit motives to perform abstraction during both training and test influence the degree to which transfer is observed in AGL (see also Gomez, 1997).

In the next section I will review some additional studies that are relevant for the basis of transfer and the formation of surface-independent representations in AGL. These studies go somewhat beyond the standard transfer paradigm (i.e., observe sequences and then classify new ones in a new domain) in various ways and in various degrees. In addition, some of the studies are not directly framed in terms of AGL, but are still similar enough to deserve to be discussed, in order to shed light on whether they may have implications for the basis of transfer in AGL.

Beyond the Standard Transfer Paradigm

Some studies have found what may seem to be particularly strong evidence for incidental formation of representations that allow for generalization to new symbol sets in AGL studies. In these studies, explicit motives to perform abstraction at test have been minimized by using indirect test tasks where there is no mentioning of underlying regularities until after the experiment is over (A.S. Reber, 1969; see also Kuhn & Dienes, 2005) or by testing infants, who presumably have no explicit expectations and motives to engage in strategic abstraction processes (Gomez & Gerken, 1999; Marcus et al., 1999; Marcus et al., 2007).

As already discussed, Newell and Bright (2001) found no transfer to new symbol sets in AGL using preference judgments as an indirect test task. In contrast, A.S. Reber (1969) found significant transfer to new symbol sets using recall tests during the learning phase as an indirect index of sensitivity to structural regularities, suggesting that surface-independent representations may form even though the participants are never told during the experiment about the presence of underlying regularities. However, Redington and Chater (2002) reported two unsuccessful attempts to replicate A.S. Reber´s (1969) results. In effect, it is not clear how replicable the finding of A.S. Reber (1969) really is. There are also alternative explanations for the transfer effect found by A.S. Reber (1969). As discussed by Jamieson and Mewhort (2005), if participants develop a set of retrieval strategies for one domain, then of course the same set of retrieval strategies will be effective in a new domain where new stimulus forms are shown and where the underlying structure of the sequences are the same in in both domains. This does not entail that the underlying rules of the sequences are abstracted into a surface-independent form. Instead, it may simply indicate that the participants have developed efficient forms of chunking and organizing a sequence when attempting to memorizing it.

Marcus et al. (1999) found that infants, for which strategic explicit processing is presumably minimal, could generalize simple grammars to new symbol sets. However, various associative and perceptual explanations of the results of Marcus et al. (1999) have been suggested that do not involve automatic extraction of algebraic surface-independent rules (e.g. Christiansen & Curtin, 1999; McClelland & Plaut, 1999; Perruchet & Vinter, 2002; Seidenberg & Elman, 1999). For example, it seems that quite elementary perceptual and/or motoric coding of same-different relationships can account for these results (e.g., coding the sequence MXX as “different-same” at the perceptual level enables efficient later processing of the sequence VTT using the same coding scheme). Much work remains to be done in empirically teasing apart different explanations in this context. Nevertheless, currently it is fair to say that the data do not warrant an explanation in terms of automatic extraction of algebraic surface-independent rules.

Conway and Christiansen (2006) suggested an approach different from the standard transfer paradigm in order to test the formation of surface-independent representations in AGL. In their Experiment 1, each participant was exposed to regularities from two different grammars during training. The presentation of regularities was similar to that used in most statistical learning studies in that one symbol was shown at a time, rather than an entire sequence being showed at a time, which is more common in AGL studies. The two grammars were instantiated in different modalities (tones vs. colors). At test, participants were asked to classify new sequences with respect to the underlying rules. Unbeknownst to the participants all of the test sequences were grammatical (half from each grammar). Half of the participants received all test sequences instantiated in one of the domains (e.g., tones) and half of the participants received them instantiated in the other domain (e.g., colors).

In the experimental procedure of Conway and Christiansen (2006), accounts that suggest surface-independent representations predict that participants should classify sequences as equally grammatical regardless of whether the domain for a particular grammar has changed between training and test. For example, if all sequences are received as tones at test and half of these belong to grammar 1 that was instantiated as colors during training and half belong to grammar 2 that was instantiated as tones during training, then test sequences from grammar 1 and 2 should be endorsed to the same extent. On the other hand, accounts that propose surface-dependent representations predict that domain preservation should increase the tendency to endorse sequences, so that sequences from grammar 2 should be endorsed more than sequences from grammar 1 in the example above. Conway and Christiansen (2006) found support for surface-dependent representations.

Using a statistical learning paradigm Turk-Browne, Junge, and Scholl (2005) found that knowledge of regularities embodied in a long stream of shapes was unaffected by changing the color of the shapes at test, indicating abstraction away from surface features. In their procedure, the stream of shapes was made of two interleaved separate streams of different shapes. The two streams were made of different colors, in order to implement an attentional manipulation so that participants mainly focused on one of the streams. At test, the participants were given a recognition test for parts of the stream that occurred frequently and parts that occurred only rarely, but all these test sequences were shown in black. Participants performed above chance for the attended stream but at chance for the unattended stream. However, a more compelling demonstration to demonstrate surface-independent abstraction of the kind discussed in this thesis would be if sequence knowledge was unaffected by switching both shapes and colors at test (although such an implementation may be difficult for the statistical structures used in these statistical learning experiments). In terms of abstraction, the results of Turk-Browne et al. (2005) demonstrate important constraints with respect to which perceptual features are necessary in statistical learning situations (color reinstatement is not necessary when participants are cued by shapes), but hardly the kind of algebraic surface-independent abstraction (e.g., Manza & A.S. Reber, 1997; Marcus et al., 1999) that has been (and is) controversial in AGL. For example, an AGL analogue of the abstraction demonstrated by Turk-Browne et al. (2005) would be if participants were trained on black letter sequences and then were able to transfer that knowledge to test sequences instantiated in the exact same letters but written in blue. Above-chance performance in such an experiment does not entail abstraction to any larger degree, since the letters are preserved and provide a powerful surface cue.

Goschke and Bolte (2007) found evidence for a kind of abstraction in a serial naming task where participants named objects (e.g., *horse*) from different categories (e.g., *animal*) sequentially. The sequence was structured in terms of the categories of the objects, but random in terms of the sequence of specific objects and overt responses. Response time analyses showed that the participants adapted to the sequential category structure even without noticing the structure. Again, however, the kind of abstraction implied by these results (although interesting in many ways) are not at the core of the debate regarding surface-independence in AGL. As noted by the authors themselves, the mapping between specific objects and the associated category is already completely in the participants´ memories before they enter the experiment (e.g., a *horse* being an *animal*). In effect, the conceptual abstraction associated with the serial naming task used by Goschke and Bolte (2007) is a by-product of semantic priming rather than the result of an algebraic abstraction mechanism.

Manza and A.S. Reber (1997) suggested that surface-independent representations may not be formed to an exhaustive extent because of the rather short training phases used in many AGL studies. The authors proposed that as learning proceeds an abstractor mechanism becomes increasingly more active enabling generalization to new surface formats. This would explain the transfer decrement phenomenon discussed earlier, that is, the fact that classification in a new domain is generally worse than same-domain performance. However, Pacton, Perruchet, Fayol, and Cleeremans (2001) investigated children´s sensitivity to real-life orthographic regularities that are not explicitly taught and found a consistent transfer decrement in different age groups that had been differentially exposed to these regularities over several years, suggesting formation of surface-dependent representations.

Study III in this thesis used the procedure developed by Conway and Christiansen (2006) to further investigate the impact of extended exposure on the development of surface-independent vs. surface-dependent representations in AGL. Study III is important because it provides a direct test of the idea that extended incidental exposure promotes abstraction toward surface-independence (e.g., Manza & A.S. Reber, 1997).

On the Nature of Abstraction Across Surface Form in AGL: Summary

The question of the extent to which AGL and learning of regularities generally is associated with surface-independence, for example, in the form of algebraic rules, is an important topic that has not been entirely resolved and is actively debated (e.g., Bonatti, Pena, Nespor, & Mehler, 2006; Pena, Bonatti, Nespor, & Mehler, 2002; Perruchet, Tyler, Galland, & Peereman. 2004; Perruchet, Peereman, & Tyler, 2006). The evidence from most AGL studies suggest that surface-dependent representations formed during training in combination with abstraction at test can account for most of the transfer studies. To the extent that surface-independent representations are formed during incidental training, they are plausibly directly or indirectly associated with tasks and processing strategies that encourage such abstraction (Redington & Chater, 2002), an explanation which is very much in line with the episodic-processing account of implicit learning (Whittlesea & Dorken, 1993). In addition, direct abstract relational coding may plausibly occur at the perceptual encoding level (Perruchet & Vinter, 2002).

Additional studies of transfer in AGL are very much needed, especially concerning the role of perceptual abstract encoding in transfer and the impact of different kinds of structural regularities and their instantiation (see Study III in this thesis for additional discussion).

The Conscious Status of the Knowledge Acquired in AGL

Apart from the question of what participants learn in AGL, there is also the question of whether the knowledge acquired and applied in AGL is conscious or not. Although conceptually separate, these two questions are linked to each other, because it is important to know what participants learn in AGL in order to settle the question of the conscious status of the knowledge (Shanks & St. John, 1994). Traditionally, different methods have been used in order to assess the conscious status of knowledge in AGL. These methods involve different assumptions of what it means for knowledge to be conscious (Gaillard, Vandenberghe, Destrebecqz, & Cleeremans, 2006). In what follows I will briefly review five different methods that have been used in AGL, namely *verbal reports*, *objective tests*, *confidence judgments*, *post-decision wagering*, and *opposition logic*. Then, I will discuss the question of conscious awareness more generally, partly in combination with a constructive/inferential framework.

Measuring the Conscious Status of Knowledge in AGL

Verbal Reports

One of the most seemingly straightforward ways of determining the conscious status of knowledge in AGL is to simply ask the participants what they can *verbally report* about the rules or how they classify sequences. A.S. Reber (1967) argued that because participants were unable to verbally describe the rules, despite classifying test sequences with above chance accuracy, the knowledge of the underlying regularities had to be unconscious. The basic procedure of verbal report has been used in many studies and participants generally have difficulties describing the underlying rules of the grammar (for reviews and discussion see Cleeremans, Destrebecqz, & Boyer, 1998; Gaillard et al., 2006; Shanks & St. John, 1994). The underlying logic of using verbal reports in this way is that of *dissociation*. This logic assumes that unconscious knowledge is demonstrated when a test of awareness (e.g., verbal report) dissociates from a test of knowledge elicitation (e.g., classification). More specifically, unconscious knowledge is demonstrated when $A = 0$, while $B > 0$, where A is a test of awareness for a piece of knowledge that is elicited by test B.

Despite the conceptual straightforwardness of verbal reports as a measure of conscious knowledge, the method is associated with a number of problems. First, verbal reports may be rather insensitive compared to other ways of eliciting knowledge, so that participants refrain from reporting knowledge held with low confidence or refrain from reporting knowledge because of vaguely formulated questions (Shanks & St. John, 1994). Second, in order for a verbal report test to be valid as a measure of conscious knowledge it is important that participants have acquired and apply the knowledge the experimenter has in mind when they classify sequences in AGL (Shanks & St. John, 1994). In effect, asking participants to report the rules of the grammar when their classification judgments may not involve knowledge of the rules of the grammar does not mean that the acquired knowledge is unconscious. Third, knowledge may be stored in a way that renders it relatively inaccessible to verbal report, but nevertheless accessible to awareness when it is applied (Shanks, 2005), for example, as in the case of motoric knowledge or conscious experiences of perceptual fluency. Fourth, the potential complexity of the knowledge acquired in AGL (Pothos, 2007) may make it difficult to report the knowledge verbally, even though the knowledge may not necessarily be unconscious.

There are other potential problems with verbal reports as well (e.g., that verbal reports may partly be a result of confabulation or operating under the wrong theory of what knowledge one has to have in order to perform a task; see Nisbett and Wilson, 1977), but the issues in the previous paragraph are enough to demonstrate why most AGL researchers no longer rely exclusively on verbal report in order to determine the conscious status of knowledge. This does not mean that verbal reports are useless. The question of what participants verbally report in AGL or any other paradigm may be an interesting research question in and of itself. Additionally, verbal reports may be used as a way of understanding how participants experience the task.

Objective Tests

Objective tests of awareness do not ask participants to report what knowledge they have, but instead require the participants to perform a direct task *A* that can be assessed with respect to some chance level. If participants demonstrate *knowledge* through their performance on a task *B* (e.g., classification of test sequences in AGL), then *A* may be said to constitute an objective measure of awareness if *A* can be assumed to require *conscious knowledge* in order to be performed at above chance levels and if *A* measures the same knowledge as that which is expressed in *B*. If participants perform above chance on both *A* and *B*, then there is evidence for conscious knowledge. If participant perform at chance on *A* but above chance on *B*, then there is possibly evidence for unconscious knowledge (if the test is maximally sensitive).

It is important to note that objective tests of awareness are fundamentally different from verbal reports (and also different from confidence judgments, discussed in the next section) in that objective tests require the participant to form a judgment with respect to states of the world, rather than with respect to mental states (Dienes, 2008). Objective tests require judgments regarding external properties (e.g., whether a sequence was shown previously in the experiment or not). Verbal reports ask participants to report on internal mental states (e.g., what knowledge a participant thinks that he or she has).

In order to illustrate objective tests of awareness, consider a study by Perruchet and Pacteau (1990, Experiment 3). In that study, participants were exposed to sequences from an artificial grammar and then they performed a recognition test on 25 different bigrams. The participants were clearly able to distinguish the bigrams that had occurred in the training sequences from the bigrams that had not occurred. Also, the correlation between the frequency of occurrence of the bigrams that had occurred in the exposure phase and the recognition scores was .61. Thus, frequency of occurrence was related to recognition scores. Furthermore, Perruchet and Pacteau then attempted to simulate classification scores of grammaticality in earlier experiments (with participants randomly drawn from the same population) by assuming that participants would classify an exemplar as ungrammatical when it contained at least one unrecognized bigram (see Perruchet and Pacteau, 1990, p. 270 for details regarding this simulation). The results showed that the simulated performance was very close to the observed performance of the participants. Based on these results, and assuming that recognition is an objective measure of awareness, the authors argued that conscious knowledge of bigrams can account for classification of artificial grammar sequences. Additionally, if conscious knowledge of very simple regularitites can account for performance, there is no need to postulate a sophisticated cognitive unconscious that performs extraction of complex rules (see Perruchet & Vinter 2002, for arguments along these lines).

Although objective tests differ from verbal reports as noted above, the underlying logic is similar in that both types of tests are usually implemented in the context of dissociation logic (or the reverse, namely association, in the case of Perruchet and Pacteau, 1990). The recognition test is assumed to test for conscious knowledge and if that conscious knowledge can account for performance on the classification test then one may argue that the knowledge responsible for the performance on the classification test is not at all unconscious (and the opposite conclusion if the objective test of awareness shows no evidence for conscious knowledge, provided that the test is sensitive enough and probes the relevant knowledge).[4]

However, there are many problems with objective tests when it comes to assessing unconscious knowledge. The most serious problem has to do with the close relation between objective tests and dissociation logic.Traditional dissociation logic requires that each of the tests trigger unique processes. That is, there is an assumption of "process-purity". This means that in order to show that knowledge is conscious or unconscious one has to have a highly sensitive direct objective test that is dependent *only* on the relevant conscious knowledge for above chance perfomance. Most researchers would probably agree that the process-purity assumption is highly unlikely (e.g., Fu et al., 2008; Higham, Vokey, & Pritchard, 2000; Jacoby, Yonelinas, & Jennings, 1997). For example, why should the recognition task used by Perruchet and Pacteau (1990) involve only conscious knowledge (given that there is both unconscious and conscious knowledge)? A.S. Reber (1997) has argued against objective tests of conscious knowledge on the grounds that they may often involve unconscious knowledge as well.

In order to avoid reliance on the process-

4 Of course, there is *no guarantee* that the fact that one can account for classification in terms of recognition scores implies that the knowledge underlying the recognition scores was actually used in classification. One could in principle just as well argue the other way around and say that the knowledge underlying the recognition scores is unconscious because it is the same as that used in the classification test, that is, unconscious knowledge of structure. In the end, it pretty much all comes down to what one takes as primary in such situations. For some the implicit is the primary (A.S. Reber, 1997) and for some the explicit is the primary (Perruchet & Vinter, 2002; Shanks & St. John, 1994). Even these issues are debatable of course, so the situation is not necessarily a dead end.

purity assumption, Reingold and Merikle (1988; Merikle, 2003) suggested a methodology involving closely matched direct and indirect tests that differ only in terms of task instructions. For example, the direct test may instruct participants to use knowledge from a previous learning phase, while the indirect test does not. Unconscious knowledge is demonstrated when the indirect test is shown to be more sensitive than the direct test, if it can be assumed that conscious knowledge, if available, would be reflected *at least* as much on the direct as on the indirect test. Using preference judgments as an indirect test and classification judgments as a direct test Kuhn and Dienes (2005) implemented such a procedure in experiments on musical rule learning using a biconditional grammar and found evidence for unconscious knowledge. However, Whittlesea and Price (2001) showed that different tests (e.g,. preference and recognition judgments) may give rise to various kinds of dissociations that can be reversed depending on the task context in which the tests are implemented. For example, preference judgments usually trigger non-analytic processing strategies, but may be induced to trigger analytic processing (and different performance) if the participants are asked to justify their judgments (Whittlesea & Price, 2001). In effect, the way knowledge representations are utilized and applied may affect which kind of knowledge becomes or does not become conscious, and there may be no default with respect to what particular kind of knowledge is triggered generally in a certain kind of test (Helman & Berry, 2003). This makes the use of objective tests of awareness problematic.

Confidence Judgments

Some researchers have argued that measures of awareness based on *confidence judgments* (also referred to as "subjective measures") may be more appropriate than objective tests in order measure the conscious status of knowledge (Dienes, Altmann, Kwan, & Goode, 1995; Dienes & Berry, 1997; Dienes & Perner, 1999). According to the logic of measures based on confidence judgments, knowledge is unconscious when confidence is unrelated to accuracy, and conscious when confidence is related to accuracy. More generally, participants have unconscious knowledge when they have objective knowledge (as revealed by a test of knowledge elicitation) that they are not subjectively aware of having (as revealed by confidence judgments).

According to the *guessing criterion*, there is some unconscious knowledge when participants perform above chance or above the level of an untrained control group even when participants report that they are guessing (i.e., they are 50 % certain of having made a correct answer). According to the *zero-correlation* criterion, there is some conscious knowledge when participants´ confidence judgments are significantly correlated with accuracy. The zero-correlation criterion is sometimes implemented as the *Chan difference score*, which is the difference between the average confidence for correct answers and the average confidence for incorrect answers (in other contexts, this measure is usually referred to as *Slope*, e.g., Yates, 1994). If participants are more confident for correct than for incorrect answers then there is some conscious knowledge in operation (for further description of these measures and some of their assumptions, see Dienes, 2008; Dienes & Perner, 1999, 2004).

According to Dienes and Perner (1999) implicit learning results in unconscious knowledge because the knowledge is not labelled as knowledge in the learning process, even though it may be used as knowledge. Accordingly, one would expect participants in artificial grammar learning to have poor metaknowledge of their knowledge according to the above and other metacognitive measures (the measures are metacognitive because they ask participants to evaluate their own knowledge). The results from AGL research have been somewhat mixed on this issue. Dienes et al. (1995) found evidence of unconscious knowledge in several experiments according to both the guessing and the zero-correlation criterion. The zero-correlation criterion revealed conscious knowledge for letter sequences consisting of only three letters, but not for strings longer than three letters, and there was more conscious knowledge when a forced-choice test procedure was used as compared to simply checking sequences from a list. Furthermore, the authors found evidence of unconscious knowledge according to the guessing criterion, and this knowledge (held with 50 % confidence) was unaffected by performing a secondary task during testing as compared to knowledge held with higher than 50 % confidence. Dienes and Scott (2005) found evidence for unconscious knowledge according to the guessing criterion, and conscious knowledge according to the zero-correlation criterion. Dienes and Altmann (1997) also found conscious knowledge according to the zero-correlation criterion, but when participants were tested in a transfer paradigm there was no evidence for conscious knowledge. Tunney and Altmann (2001) also found evidence for unconscious knowledge according to the zero-correlation criterion in a transfer paradigm (although the results sometimes approached significance toward showing conscious knowledge).

Allwood, Granhag, and H. Johansson (2000, Experiment 2) found that participants exhibited quite good metaknowledge according to calibration measures. Furthermore, Tunney and Shanks (2003a) found that participants had conscious knowledge according to a version of the zero-correlation criterion based on binary confidence judgments (high/low confidence) and a signal detection analysis of metaknowledge. The authors concluded that participants, for some unknown reason, find it easier to place their phenomenal states on a binary rather than a continuous (50-100 %) confidence scale. Tunney (2005) extended these results by showing that participants´ binary confidence judgments were more sensitive than continuous confidence judgements to various measures of similarity in AGL (e.g., item-specific similarity and bigram chunk strength). Using a biconditional grammar, Channon, Shanks, Johnstone, Vakili, Chin, and Sinclair (2002) found evidence for unconscious knowledge according to both the zero-correlation and guessing criteria for amnesic and control participants.

Although the results are mixed overall, AGL studies typically show evidence for some unconscious knowledge according to the guessing criterion and for some conscious knowledge according to the zero-correlation criterion (Dienes & Scott, 2005).

A potential problem with some confidence judgment measures of the conscious status of knowledge may be that it is up to the participants to set their own ”confidence criterions”, for example, what kind of internal state should count as a guess (Tunney & Shanks, 2003a). More generally, it may also be the case that the confidence associated with particular judgments is not only a result of the conscious accessibility of the underlying knowledge, but also a result of many other factors, such as the processing strategies adopted by participants (Whittlesea, Brooks, & Westcott, 1994) and the inferences and expectations associated with assessing one´s own knowledge (Schwartz, Benjamin, & Bjork, 1997; Whittlesea & Dorken, 1997). Furthermore, as pointed out by Higham et al. (2000) measures based on confidence judgments are also based on dissociation logic in the sense that unconscious knowledge is demonstrated when one test (based on confidence judgments) dissociates from another test (e.g., classification judgments).

Post-Decision Wagering

Persaud, McLeod, and Cowey (2007) introduced a measure that they referred to as *post-decision wagering* in order to measure the conscious status of knowledge in a variety of situations (e.g., blindsight, the IOWA gambling task, and AGL). Post-decision wagering involves making a decision and then placing a wager (high or low) on whether the decision is correct or not. Correct decisions are rewarded with the wagered amount and for incorrect decisions the wagered amount is deducted from one´s earnings. The objective is to earn as much money as possible, a task that requires strategic wagering according to what knowledge one actually has. The ultimate outcome, of course, is to bet high on all correct decisions and low on all incorrect decisions.

Persaud et al. (2007) argued that post-decision wagering is a more objective and direct measure of awareness then confidence judgments. According to the authors the latter involves introspecting about one´s own knowledge, while the former involves no, or at least less, introspection. Participants in Persaud et al. (2007) described the post-decision wagering task as intuitive and fun, suggesting that the task may be more natural than making confidence judgments.

Persaud et al. (2007) found evidence for both conscious and unconscious knowledge in AGL using post-decision wagering and real money. For example, when the participants classified correctly they did not wager high more often than 45 % of the times (indicating a lack of conscious knowledge), and 77 % of the low wagers were made after correct decisions (indicating many wasted opportunities to make money). However, a high wager was more likely after a correct decision (45 % high wagers) than after incorrect decisions (32 % high wagers), indicating some conscious knowledge. Furthermore, the authors also tested participants who wagered imaginary money. This group made less accurate classification decisions than the real money group, but there was no difference in wagering performance.

It is too early to say anything general about post-decision wagering, since the method has not been applied to any larger extent yet. The method seems promising, but the alleged directness of the measure has also been questioned (Seth, in press) and the metacognitive components of post-decision wagering makes it potentially susceptible to the criticism directed at measures based on confidence judgments. For example, since post-decision wagering involves a participant making a decision (wagering) on the basis of the judged accuracy of a previous decision, surely post-decision wagering has to involve some kind of attempted monitoring of the mental states that went into the previous decision. As mentioned previously in the section on *Confidence Judgments*, all kinds of factors presumably enter into that kind of monitoring in addition to the conscious accessibility of the underlying knowledge.

Opposition Logic

In *opposition logic* (Jacoby, Woloshyn, & Kelley, 1989) conscious and unconscious influences are conceptualized in terms of control, in the sense that conscious influences are assumed to be subject to strategic control and unconscious influences are assumed to be beyond strategic control. These two influences are pitted against each other in an experimental context, rather than being directly estimated by separate tests (as in dissociation logic). It is assumed that both conscious and unconscious influences typically affect performance on any task (Jacoby, Yonelinas, & Jennings, 1997). The implementation of opposition logic involves two crucial conditions, *inclusion* and *exclusion*. In the inclusion condition participants perform a task where conscious and unconscious influences act in concert and in the exclusion condition the task is designed so that conscious and unconscious influences act in opposition. Conscious (controllable) knowledge is demonstrated if participants in the exclusion condition can refrain from applying knowledge so that performance is significantly below the inclusion group. Unconscious (uncontrollable) knowledge may simultaneously be demonstrated if the exclusion group nevertheless performs above some appropriate baseline level despite being instructed to suppress expression of knowledge. Opposition logic has been used in order to investigate the presence and properties of unconscious knowledge in the SRT task (Destrebecqz & Cleeremans, 2001; Fu et al., 2008; Wilkinson & Shanks, 2004), but here I will focus on the implementation of opposition logic in AGL.

Higham et al. (2000, Experiment 1) exposed participants to sequences from two grammars (A and B). At test, participants classified new sequences from grammar A (GA) and grammar B (GB), and also sequences ungrammatical (U) with respect to both grammars. Participants in the inclusion condition were instructed to endorse GA and GB sequences as grammatical, while participants in the exclusion condition were instructed to endorse only GB sequences. The authors found evidence for conscious influences in the sense that the endorsement rate of GA sequences was higher in the inclusion than in the exclusion group. Thus, participants in the exclusion group were to some extent able to avoid endorsing GA sequences, despite having been exposed to other sequences from the A grammar during training. At the same time, the authors argued that there was evidence for unconscious influences in the sense that, within the exclusion condition, the participants endorsed GA sequences at a higher rate than they endorsed U sequences, despite not intending to do so. Furthermore, invoking a response deadline reduced conscious influences but left unconscious influences unaffected.

Dienes et al. (1995) used a similar implementation of opposition logic but found no evidence for unconscious influences. In their experiments, the participants were able to selectively apply knowledge restricted to one of the two grammars. The reason for this is most likely that the authors used two grammars that were more distinct from each other than the two grammars used by Higham et al. (2000).

Opposition logic embodies many potential virtues. It does not rely on verbal report, is not directly derived from dissociation logic, and does not rely on the process-purity assumption (that different tasks selectively trigger specific processes). However, despite the potential virtues some concerns have been raised about opposition logic, both generally and with respect to the specific implementation within AGL.

Some researchers have argued against opposition logic as a measure of the conscious status of knowledge because it is not really about conscious awareness, but rather about control, which is conceptually orthogonal to awareness (Redington, 2000; Perruchet & Vinter, 2002). The basic idea is that knowledge may be difficult to control despite being conscious to a large extent. Furthermore, the basic results regarding conscious and unconscious influences are often validated through obtaining dissociations of various kinds (e.g., by invoking a response deadline, as in Higham et al., 2000), thus reintroducing dissociation logic one step further down the path and not really avoiding it. Also, since opposition logic *assumes* separate influences the procedure is biased against single-process theories which do not assume separate processes. In fact, opposition logic may incorrectly indicate that data generated by a single process should be partitioned so as to having been derived from separate processes (Ratcliff, McKoon, & Van Zandt, 1995).

More specific to AGL, Redington (2000)

argued that participants in the exclusion condition in Higham et al. (2000) may have incorrectly endorsed GA sequences because of the high similarity between these sequences and the GB sequences, not because of unconscious knowledge of regularities. Tunney and Shanks (2003b) provided connectionist simulations of the experiments in Higham et al. (2000) using a simple recurrent network model that does not embody separate processes (controlled and automatic). The model was able to reproduce the basic results of Higham et al., suggesting that a single similarity-based account may explain the results. However, Vokey and Higham (2004) argued that the simulations in Tunney and Shanks (2003b) did not provide clean fits to the dissociations obtained by Higham et al. (2000), and Vokey and Higham presented their own simulations using an autoassociative network model and maintained that the data of Higham et al. (2000) shows evidence for separable influences in AGL (controlled and automatic).

Part of the debate described above concerning controllability in AGL seems to be a terminological issue. There is no doubt that the participants in the experiments of Higham et al. (2000) were subject to unintentional influences in that they endorsed sequences from a grammar they did not intend to endorse. The reason for this can be traced to the similarity between the two grammars. However, a further question is whether the results necessarily reflect a separate uncontrollable computational process. The simulations of both Vokey and Higham (2004) and Tunney and Shanks (2003) suggest that the data from Higham et al. (2000) do not necessitate a separate uncontrollable process.

Study IV in this thesis used an implementation of opposition logic in AGL different from the one used by Higham et al. (2000). In the procedure of Study IV, participants were trained on only one grammar and were then required to generate sequences under either inclusion or exclusion instructions (this kind of implementation has been used in the SRT task, e.g., Destrebecqz & Cleeremans, 2001). This procedure is arguably more straightforward than the procedure in Higham et al. (2000) and not directly susceptible to criticism based on similarity between different grammars (e.g., Redington, 2000).

A Constructive Inferential Perspective on Awareness

Whittlesea and Dorken (1997) argued that “implicit learning is just ordinary learning without becoming aware of the implications of that learning for performing unanticipated activities” (p. 63). As it stands, the quoted passage may sound like a characterization of implicit learning partly in terms of awareness, but that is not the authors´ intention. In their article, Whittlesea and Dorken argue that no conception or measure of awareness supports the notion of two kinds of representationally distinct knowledge bases (e.g., implicit vs explicit). The representations that drive performance on any occasion are all unconscious so that people have no direct conscious access to them. Furthermore, when people evaluate what they know, when they say something about their abilities, then they rely on intuitive theories and expectations or whatever relevant information is available in the current context in order to make sense of their performance. In other words, awareness is an inferential activity, rather than mere activation of knowledge.

Instead of two knowledge bases, there may instead exist two or more uses of knowledge. That is, one and the same knowledge base may be accessed in different ways and create differences in performance and subjective experience (Kinder et al., 2003; Whittlesea, 2003; Whittlesea & Leboe, 2000). The confidence with which people perform a task is thus not a matter of having access to explicit or implicit knowledge, but rather a matter of (largely unconscious) evaluation and inference.

People do not always know the distal reasons for and the inferences behind why they choose the way they do (Nisbett & Wilson, 1977), why they settle on a certain degree of confidence (Whittlesea, Brooks, & Westcott, 1994), why they evaluate their own knowledge in particular ways (Sanna & Schwartz, 2007; Schwartz, Benjamin, & Bjork, 1997; Werth & Strack, 2003), why they (sometimes erroneously) justify their preference choices in certain ways (P. Johansson, Hall, Sikström, & Olsson, 2005), or why they classify exemplars in a certain way (Wright & Whittlesea, 1998), but they may form all kinds of ideas as to why they do behave the way they do and this may sometimes appear as pre-computed knowledge even if it is not. The experiment by Wright and Whittlesea (1998, Experiment 5) described earlier serves as an illustrating example of the constructive nature of awareness.

According to the constructive and inferential perspective on awareness offered by Whittlesea and Dorken (1997) what people become aware of is determined by the construction and realization of knowledge out of a knowledge base that exists as a potential. This perspective does not preclude the study of awareness in implicit learning, but it renders the alleged dissociation between different learning modes and awareness problematic. An important issue that emerges from the constructive inferential perspective is how different conditions during the test phase influence the application of knowledge and what kind and degree of awareness is realized based on evaluation of that knowledge application (for examples in the memory literature, see Goldinger & Hansen, 2005; Leboe & Whittlesea, 2002; Whittlesea, 2002, 2004).

With this constructive and inferential perspective in mind, Study I and II investigated the effect of non-analytic processing on awareness of knowledge in AGL as assessed by confidence judgments. Non-analytic processing was manipulated either by speeded responding (Study I) or by instructions (Study II). It was expected that non-analytic processing might actually lead to better meta-knowledge in AGL, since non-analytic processing can be expected to make the participants fall back on the resemblance heuristic and/or the fluency heuristic to a greater extent than normal, and rely on their general impression of a test item. The metacognitive feelings that accompany processing may sometimes be more veridical under conditions of non-analytic processing compared to analytic processing, at least in learning environments where it is difficult to verbalize the knowledge and when the participants do not themselves expect to perform well (as is typical in AGL tasks).

The conscious status of the knowledge acquired in AGL: Summary

Currently, there is no direct consensus concerning the conscious status of the knowledge acquired in AGL. A variety of different methods have been used in order to assess the conscious status of knowledge in AGL, and each of these methods have their own particular assumptions, which are often debated and questioned. It is quite clear that unconscious knowledge is by no means a ubiquitous feature of implicit learning, but knowledge may still be applied without awareness under various circumstances. Rather than trying to relate conscious vs. unconscious knowledge to particular kinds of learning or particular kinds of representations, it may be more fruitful to investigate the processes and attributions underlying different kinds of subjective phenomenology in AGL. Study I and II in the present thesis attempted to do just that, by investigating the relation between analytic/non-analytic processing and awareness based on confidence judgments in AGL.

Purpose of the Studies

The general purpose of all the studies included in this thesis was to investigate critical issues concerning the nature and conscious status of the knowledge acquired in AGL. The studies are quite independent from each other and, for the most part, touch on what may be viewed as rather separate issues within AGL (Study I and II are partly related though). Even though the studies can be viewed quite separately, the *General Discussion* section offers elaborations on possible ways of integrating the results.

Study I investigated the use of a perceptual fluency heuristic when making grammaticality judgments in AGL in combination with non-analytic processing at test. Fluency is often talked about as a potential cue for sensitivity to structure in AGL, but has rarely been directly manipulated and investigated in AGL (for an exception, see Kinder et al., 2003).

Both Study I and Study II investigated the effect of non-analytic processing at test on metacognitive discrimination through confidence judgments in AGL. It was expected that non-analytic processing might increase metacognitive discrimination as a result of the participants´ relying more on feelings of familiarity or fluency when judging their own knowledge.

Study III investigated the controversial issue of abstraction toward surface-independence in AGL and provided a direct test of the hypothesis that extended exposure is associated with increased abstraction (Manza & Reber, 1997). This hypothesis has, to my knowledge, never been tested before in AGL within a paradigm where surface-dependent and surface-independent representations are pitted against each other.

Study IV used a novel implementation of opposition logic in AGL using a generation task in order to investigate the controllability of knowledge in AGL, an issue that has been somewhat controversial (Dienes et al., 1995; Higham et al., 2000; Tunney & Shanks, 2003b; Vokey & Higham, 2004). One part of the controversy relates to the proper way of implementing opposition logic and how the results may depend on factors that are unrelated to controllability (Redington, 2000). The procedure used in Study IV is arguably more straightforward than previous implementations (e.g., Higham et al., 2000) and is similar to the implementation often used in the SRT task (Fu et al., 2008).

Overview of the Studies

In the sections that follow, the studies included in the present thesis (Study I, II, III, and IV) are briefly described, discussed, and related to questions regarding knowledge representation and awareness in AGL. This is then followed by a more general discussion taking a slightly wider perspective, including some notes with respect to future research. The studies are described in much more detail in the actual manuscripts, all of which are included as the final part of this thesis.

Study I

The main aim of Study I was to investigate the relation between non-analytic processing and the use of a fluency heuristic in AGL. Kinder et al. (2003) used a perceptual clarification procedure in order to manipulate fluency in the test phase of several AGL experiments and found that participants relied on a fluency heuristic when making classification judgments (and also under some conditions when making recognition judgments). These results are consistent with the idea that fluency may be used as a cue when making judgments about the structure of a domain, especially since grammatical test sequences are processed more fluently than ungrammatical test sequences after exposure to grammatical training sequences (Buchner, 1994).

In Study I, perceptual fluency was manipulated by way of *masked priming*, a manipulation that has been shown to affect recognition judgments (Jacoby & Whitehouse, 1989; Kurilla & Westerman, 2008) and preference judgments (R. Reber, Winkielman, & Schwarz, 1998). In the test phase, before the presentation of a to-be-classified target item, the target item was flashed briefly on screen in between so called “masks” (arbitrary symbols intended to make perception of the flashed item difficult) on half of the trials. If the target item has been flashed briefly just before it is presented, this may increase the fluency with which it is processed once it appears on screen. If the participants rely on a fluency heuristic then they should endorse target items that have been primed (i.e., where the target item has been flashed briefly) more than those that have not been primed.

It has been suggested that non-analytic processing is related to the use of a fluency heuristic in various paradigms (Kinder et al., 2003; Whittlesea & Price, 2001). In Study I, non-analytic processing was encouraged by invoking response time restrictions. In Experiment 1 there were no response time restrictions, while in Experiment 2 the participants had to respond within 2000 ms. Experiment 3 used a response-signal procedure requiring participants to respond within a time window of 500 ms after delays of 500 ms or 2000 ms (manipulated within participants).

Study I also investigated the relation between non-analytic processing and different sources of confidence in AGL. After each classification judgment the participants indicated their confidence in having answered correctly on a scale from 50 to 100 % certainty. Tunney (2005) found that binary (high/low confidence) were sensitive to both item-specific similarity and grammaticality, while continuous confidence judgments were not sensitive to item-specific similarity. In line with Tunney (2005), Study I was intended to investigate the impact of different sources of confidence (non-specific similarity, item-specific similarity, and masked priming). Of particular interest was the question of whether non-analytic processing may affect the participants´ metaknowledge. Non-analytic processing was expected to increase the impact of non-specific similarity, either through a fluency or a resemblance heuristic (see the *Heuristics in AGL* section), since non-analytic processing usually entails relying on one´s general impression of an item. Such a result would indicate that different processing strategies may affect the phenomenology with which classification decisions are made in AGL, in line with a constructive inferential perspective on awareness.

The main result of Study I was that fluency (masked priming) had an impact on classification only under relatively non-analytic processing conditions. Specifically, masked priming increased endorsement rates in Experiment 2 (mean response time (*RT*) = 1310 ms) and at the short delay in Experiment 3 (mean *RT* = 838 ms), but not in Experiment 1 (mean *RT* = 5950 ms) or at the long delay in Experiment 3 (mean *RT* = 2280 ms).

The participants showed metaknowledge of both grammaticality and item-specific similarity according to the zero-correlation criterion (implemented as in Tunney, 2005; see Study I for details) both in Experiment 1 and 2. However, fluency manipulated through masked priming did not affect confidence to any larger extent. The participants had better metaknowledge of grammaticality under a response deadline (Experiment 2) than when tested under no deadline (Experiment 1). Confidence judgments were not analyzed in Experiment 3 because of lack of data points.

Taken together, the results of Study I support the notion that different heuristics are available to participants in AGL and classification generally (Kinder et al., 2003; Whittlesea & Leboe, 2000), the fluency heuristic being one of them. In effect, sensitivity to structure may be partly realized through reliance on fluency caused by (specific of non-specific) similarity to previous instances. Specifically, the results support the idea that the fluency heuristic may be particularly associated with conditions encouraging non-analytic processing strategies (Whittlesea & Price, 2001). Furthermore, the results showed that metaknowledge of grammaticality was better under a response deadline compared to no response deadline. Thus, non-analytic processing and heuristics may support awareness of grammaticality (or non-specific similarity) when making classification decisions in AGL, in line with an inferential perspective on awareness.

Study II

Study II was designed to investigate the relation between metaknowledge and analytic vs. non-analytic processing in AGL. Participants in AGL experiments can often distinguish between their own correct and incorrect classification decisions to some degree (Dienes & Scott, 2005; Tunney, 2005; Tunney & Shanks, 2003a), indicating awareness of the correctness of decisions. In the context of the SRT task, Norman, Price, and Duff (2006) suggested that participants may be differentially sensitive to non-focal information present in “fringe consciousness”, that is, the peripheral contents of consciousness that provide a background for more focal contents. An example of fringe consciousness may be a feeling of “rightness” that accompanies a decision. If metaknowledge in AGL is dependent on non-focal contents in fringe consciousness, then one might expect that non-analytic responding gives rise to better metaknowledge.

Non-analytic processing refers to relying on one´s general impression of a stimulus, while analytic processing is more about focusing on specific details (Whittlesea & Price, 2001). Non-analytic decisions often appear more intuitive and feeling-based than analytic decisions, in line with the former being related to contents in fringe consciousness to a larger extent than the latter. To illustrate, a person deciding to buy a music CD may offer a rather non-analytic justification, such as “I don´t know what it is, but it just sounds beautiful, the way it is composed, how it all comes together, it´s just perfect, I love it.”. Another person may offer a more analytic justification, such as “This is just what I looked for, I like how the songs and the production make room for the accordion, it is not used very often nowadays, I love that sound. I also like how the percussion provides a solid background for the harmonies.”. In the non-analytic case the decision is based on an overall feeling and in the analytic case the decision is based to a larger extent on an analysis of details and relations between them.

In order to investigate the hypothesis that non-analytic processing can support metaknowledge in AGL, a rather simple AGL task was used. After a standard exposure phase with grammatical sequences, each test trial involved one grammatical and one ungrammatical sequence. The participants´ task on each trial was to say which sequence was grammatical and then report how confident they were in their decision (from 50 to 100 %). All ungrammatical sequences contained one bigram that never appeared during the training phase. In the *analytic* condition, the participants were instructed to classify on the basis of details they could remember from the training phase, and in the *non-analytic* condition they were instructed to respond on the basis of which sequence "felt right" as a whole, without scanning the sequences for details.

The results showed that the analytic and non-analytic conditions did not differ significantly with respect to classification accuracy nor with respect to overall confidence. However, the difference in confidence between correct and incorrect classification decisions was larger (more positive) in the non-analytic compared to the analytic condition. Thus, non-analytic processing was associated with an increased ability to distinguish between correct and incorrect choices, consistent with the idea that metaknowledge and awareness in AGL may be associated with non-analytic processes and contents in fringe consciousness.

The results of Study II are consistent with other research in different implicit learning paradigms pointing to a role for non-analytic processes, such as fluency and familarity (Helman & Berry, 2003; Lotz & Kinder, 2006a; Kinder, et al., 2003; Newell & Bright, 2001; Norman et al., 2006). The results also open up interesting avenues for future research. For example, Tunney and Shanks (2003a) found that binary confidence judgments (high/low) were more sensitive than continuous ones (50-100 %). The results of Study II suggest the potentially testable prediction that, relative to continuous confidence judgments, binary confidence judgments may trigger a more non-analytic kind of introspection.

Study III

Study III set out to investigate the formation of abstract representations, in the sense of *surface-independent* representations, in AGL. Participants can transfer knowledge between different domains in AGL (e.g., Altmann et al., 1995), a finding that some researchers have taken to imply automatic formation of abstract surface-independent representations during exposure to structural regularities (Manza & A.S. Reber, 1997; Marcus et al., 1999; A.S. Reber, 1969, 1989). Other researchers have offered explanations, formal models, and data consistent with the formation of stimulus-specific representations during training (Brooks & Vokey, 1991; Christiansen & Curtin, 1999; Dienes, Altmann, & Gao, 1999; Redington & Chater, 1996; Tunney & Altmann, 2001). In addition, the finding that same-domain performance is better than transfer performance (the so called transfer decrement effect) suggests stimulus-specific representations (Pacton et al., 2001; Perruchet & Vinter, 2002).

Manza and A. S. Reber (1997; see also Meulemans and Van der Linden, 1997) argued that the reason for transfer decrement is that surface-independent representations have not had time to form fully. The authors proposed that as learning proceeds, an abstractor mechanism gradually becomes involved in the learning process. If this is the case, then extended training should result in more surface-independence in AGL. Study III investigated this hypothesis in three experiments using a modified AGL paradigm introduced by Conway and Christiansen (2006; see the *Accounting for Transfer without Automatic Abstraction during Training* section above for a description of the basic procedure) intended to pit surface-dependent and surface-independent representations against each other.

In all experiments, short (6 blocks) and long (18 blocks) training phases were compared. In Experiment 1 participants were exposed to visual sequences from one grammar and auditory sequences from another grammar. The sequences were created from the grammars shown in Figure 1 by substituting the letters in the grammars for specific colors or tones. At test, the participants classified new sequences from both grammars instantiated only as visual sequences or only as auditory sequences (half of the participants in each case). The sequences were always presented one element at a time both during training and test.

Experiment 2 served as a control experiment for Experiment 1, and the participants were only exposed to a single grammar in the learning phase. Half of the participants were exposed to tones and half to colors. At test, the participants classified sequences from the exposure grammar and another grammar. The test modality was always the same as the exposure modality. Experiment 2 thus provided a baseline in order to assess the degree of learning independently of whether the representations are surface-independent or surface-dependent.

Experiment 3 used a dual grammar design similar to Experiment 1, but only visual sequences were used (shapes for one grammar and colors for the other grammar) and all elements of a sequence were shown simultaneously (as is more common in AGL studies).

If extended training results in more surface-independent knowledge, then participants in the long exposure condition in Experiment 1 and 3 should be less inclined, compared to the short exposure condition, to classify sequences on the basis of modality correspondence between training and test. For example, say that a participant is exposed to visual sequences from grammar 1 and auditory sequences from grammar 2 during training, and then receives only new auditory sequences from both grammars at test. In this case, extended training should not increase the endorsement rates for grammar 2 sequences (preserved modality) relative to grammar 1 sequences (changed modality) if extended training is associated with surface-independence. On the other hand, such an increase in endorsement rates is exactly what one might expect if extended training gives rise to more surface-dependent knowledge.

The results of Experiment 1 and 3 showed that extended training was associated with an increased tendency to classify on the basis of preservation of instantiation of surface properties between training and test. In addition, Experiment 2 showed that the increase in *surface-dependent* knowledge associated with the transition from short to long exposure in Experiment 1 practically accounted for the entire learning advantage associated with the short-to-long transition. In effect, there was practically no evidence for the formation of abstract representations but clear and direct evidence for surface-dependent representations. The results are in line with implicit learning theories and models associated primarily with surface-dependent knowledge (Brooks, & Vokey, 1991; Christiansen & Curtin, 1999; Dienes, Altmann, & Gao, 1999; Johnstone & Shanks, 2001; Pacton et al., 2001; Redington & Chater, 1996; Whittlesea & Dorken, 1993), but not in line with those that involve automatic abstraction of surface-independent rules in AGL (Knowlton & Squire, 1994, 1996; Manza & A.S. Reber, 1997; Marcus et al., 1999; Meulemans & Van der Linden, 1997; A.S. Reber, 1969, 1989).

Some studies appear to demonstrate dissociations between abstract rules and surface-dependent knowledge in AGL (Chang & Knowlton, 2004; Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004). However, researchers have yet to specify what those abstract rules consist in (Johnstone & Shanks, 1999). Repetition patterns are an essential component of transfer in AGL (Brooks & Vokey, 1991; Gomez, Gerken, & Schvaneveldt, 2000; Lotz & Kinder, 2006b; for related experiments, see Endress, Dehaene-Lambertz, & Mehler, 2007) and it is not obvious that sensitivity to such patterns must reflect abstracted symbolic rules. Rather, repetition patterns may be learned at a much more perceptual and/or motoric level (for discussion, see Lotz & Kinder, 2006b; Perruchet and Vinter, 2002). Furthermore, even if AGL classification should reveal partly surface-independent knowledge, there is still the question of how much of that surface-independent knowledge is formed through intentional strategies at test and how much is formed during training (Brooks & Vokey, 1991; Vokey & Higham, 2005). In addition, it is likely that encoding processes moderate what kind of knowledge is formed (Jamieson & Mewhort, 2005; Johnstone & Shanks, 2001; Whittlesea & Dorken, 1993), which goes against the idea that abstract rule knowledge is formed *automatically* through the operation of an unconscious rule abstraction mechanism (A. S. Reber, 1989).

Future studies could look more closely at the influence of different encoding processes and their potential moderating effect concerning surface-independence in AGL, using experimental paradigms where it is possible to put surface-dependent and surface-independent representations in opposition. The results of Study III suggest that extended observational learning is associated with an increase in surface-dependent knowledge, rather than increased surface-independence.

Study IV

Study IV used opposition logic to investigate automatic vs. controllable influences in AGL. Typically, studies on automatic and unconscious influences in AGL have used dissociation logic, whereby two separate tests are used, one for assessing learning and one for assessing the unconscious or automatic status of that knowledge. In light of the problems associated with dissociation logic (see the section *Measuring the Conscious Status of Knowledge in AGL* above) Higham et al. (2000) instead recommended adopting opposition logic (Jacoby, Woloshyn, & Kelley, 1989) to AGL.

In opposition logic, influences that support intentional control are put in opposition to influences that do not support intentional control. Higham et al. (2000; cf. Dienes et al., 1995) trained participants on two grammars (A and B) in different contexts and then let participants classify new grammatical (from each of the two grammars) and ungrammatical sequences. At test, participants in the *inclusion* condition were asked to endorse sequences consistent with grammar A or B, while participants in the *exclusion* condition were asked to endorse only sequences consistent with grammar B. Higham et al. argued that controllable influences would be revealed by a higher endorsement rate for grammar A sequences in the inclusion than in the exclusion condition, and automatic influences by a higher endorsement rate in the exclusion condition for grammar A sequences than for sequences illegal with respect to both grammars. The authors found evidence for both types of influences.

The two-grammar designed adopted by Higham et al. (2000) was subsequently criticized (Redington, 2000; Tunney & Shanks, 2003b; see Vokey & Higham, 2004, for a response). The higher endorsement of grammar A sequences relative to ungrammatical sequences in the exclusion condition may have been due to similarity between the two grammars, rather than unconscious automatic influences.

Study IV used a design more straightforward than the one adopted by Higham et al.

(2000). In two experiments, the participants memorized sequences from an artificial grammar and were then informed of the presence of a rule structure (but not its nature). The participants were then told to *generate* new grammatical sequences (inclusion condition) or ungrammatical sequences (exclusion condition). With this procedure automatic application of knowledge is indicated by above-chance generation of sequential regularities in the exclusion condition (for a similar implementation in the context of the SRT task, see Destrebecqz and Cleeremans, 2001).

Experiment 1 investigated generation performance in both inclusion and exclusion conditions using a grammar from Jamieson and Mewhort (2005) specifying first-order dependencies (transitions depending on one element of context, e.g., C may occur after B, but not after D or E). The grammar is shown in Table 1. Base-line generation performance was assessed by training a separate group of participants on a pseudo-random grammar, which was much less constrained than the grammar in Table 1. Experiment 2 used grammars specifying second-order dependencies (transitions depending on two elements of context, e.g. A may occur after BC, but not after DC or FC), and also investigated the impact of different retention intervals on generation performance in the light of previous research suggesting that automatic influences are less susceptible to decay in AGL (Higham et al., 2000; Tunney, 2003). Instead of using a control group trained on a random grammar, Experiment 2 used a cross-over design (cf. Redington and Chater, 1996) involving two grammars that were complements of each other at the second-order dependency level (legal trigrams, i.e., subsequences of three consecutive symbols, in one grammar were illegal in the other grammar and vice versa). Half of the participants were trained on one grammar and half on the other (see Study IV for details).

Table 1. *Transitional Probabilities of the Structured Grammar used in Experiment 1 of Study IV (adapted from Jamieson and Mewhort, 2005)*

	A	B	C	D	E	F
A	0	.5	.5	0	0	0
B	0	0	.5	.5	0	0
C	0	0	0	.5	.5	0
D	0	0	0	0	.5	.5
E	.5	0	0	0	0	.5
F	.5	.5	0	0	0	0

Note. Sequences are generated by starting with one element (randomly) and then conforming to the transitional probabilities of the grammar when adding new elements. The transitional probability between element X and element Y is read out by finding the intersection between X in the left column and Y in the top row. For example, in the structured grammar, if the first element is C (left column), then the next element may be either D or E (top row, .5 probability for each). If D is chosen, then the element to follow D (left column) is either E or F (top row, .5 probability for each) and so on.

Experiment 1 showed no evidence of automatic application of knowledge in the generation task, even though the participants in the inclusion condition exhibited a small degree of sensitivity to the first-order dependencies dictated by the grammar. In Experiment 2 the participants showed no evidence of sensitivity to the dependencies of the grammar at all, regardless of retention interval.

The results of Study IV could partly be taken to suggest that AGL is not associated with automatic application of knowledge to any larger extent. However, given that there was no evidence of learning at all in Experiment 2 (as assessed by the generation task) and only a small degree of above-chance generation performance in Experiment 1, it is possible that a more sensitive type of generation task (e.g., Wilkinson & Shanks, 2004) or higher levels of learning would have been associated with detectable automatic application of knowledge. Given the results of Study IV, a question that emerges is why the participants showed such limited abilities to generate sequences containing the relevant dependencies?

The grammar used in Experiment 1 specifies that any element can start a sequence, a sequence is always six elements long, and each element is associated with straightforward transitional probabilities with respect to the next element. As shown in Table 1, if A begins a sequence, then the next element is either B or C (.5 probability for each). Say that C occurs after A in a sequence. In that case, the element that follows C is either D or E (.5 probability for each), and so on. The fact that each element is associated with first-order transitional probabilities of .5 makes it difficult to predict exactly which element will come next in a given sequence (even with full learning of the dependencies, one only has a 50 % chance of being right). However, in a free generation task scored with respect to whether the transitions within a generated sequence are grammatical or not (as in Study IV), it is not necessary to be able to predict *any particular* sequence. Since the objective is simply to generate sequences that obey the grammar *any* grammatical transitions will do.

One possibility for the low levels of generation performance in the experiments of Study IV could be that the grammars generate a lot of interference. For example, both the SRN model and PARSER are sensitive to interference in the form of prediction conflicts, such as when A may be followed by both B and C (Boucher & Dienes, 2003; Perruchet & Pacton, 2006). In contrast to the SRN, which is a statistical learning mechanism, PARSER forms chunks during learning. These chunks are sensitive to interference, so that the strength of the chunk AB is weakened by the presentation of, for example, AC (Perruchet & Pacton, 2006). In contrast to both the SRN and PARSER, the competitive chunking model (Servan-Schreiber & Anderson, 1990) is not intrinsically sensitive to interference. The competitive chunking model forms chunks during learning, but the strength of these chunks are not adjusted as a function of other chunks containing similar elements. Boucher and Dienes (2003) compared the SRN and the competitive chunking model and found that participants were not as sensitive to prediction conflicts in AGL as the SRN. Rather, the results were more in line with the competitive chunking model. If the low generation performance in Study IV were partly due to interference, then this might be reflected better by the SRN and PARSER than the competitive chunking model. Such a state of affairs would then open up for further research concerning the role of interference in AGL (Boucher & Dienes, 2003; Perruchet & Pacton, 2006).

Future studies could look closer at the role of interference in both generation and classification tasks in AGL using grammars similar to those in Study IV. When it comes to investigating automatic influences in AGL through opposition logic, the grammars in Study IV are attractive because they do not contain positional contraints (all elements can occur in all positions), making scoring of generation performance quite straightforward. In the finite-state grammars often used in AGL research there are positional constraints as well, making scoring less straightforward. Future studies could investigate whether grammars more constrained than those in Study IV can result in evidence of automatic application of knowledge. Study IV shows in a straightforward way how such an investigation can be methodologically implemented in different ways, for example, by using a control group trained on random sequences (Experiment 1) or a cross-over design (cf. Redington & Chater, 1996) involving complementary grammars.

To sum up, the results of Study IV suggest that the participants had intentional control of the little knowledge they showed evidence of having in a generation task after being trained on a grammar specifying first-order dependencies (Experiment 1). The low levels of performance suggest possible interference effects or the generation task being too insensitive. These issues could be further explored together with computational models and experimental manipulations intended to encourage automatic application of knowledge.

General Discussion

The studies reported in this thesis shed light on a number of issues regarding knowledge representation/use and the conscious status of knowledge in AGL.

Study I used a masked priming procedure to induce surprising processing fluency during parts of the test phase. The results indicated that participants use a fluency heuristic in AGL under conditions encouraging non-analytic processing, but not during more analytic processing conditions. This is in line with a heuristic view of classification in AGL, a view which emphasizes flexible and strategic processing (Kinder et al., 2003; Whittlesea & Leboe, 2000).

Study I and II both showed evidence of increased conscious judgment knowledge (the knowledge that a particular sequence is grammatical, cf. Dienes & Scott, 2005) as indicated by confidence judgements under non-analytic processing conditions. These results show that non-analytic classification decisions are not made in the absence of awareness and that the phenomenological content associated with non-analytic decisions may sometimes result in more appropriate metacognitive judgments than more analytic decisions.

Study III showed evidence of surface-dependent representations as a result of incidental observational training in AGL. Crucially, and in contrast to abstractionist accounts that emphasize surface-independence (e.g., A. S. Reber, 1969, 1989), increasing the length of the training phase did not increase the impact of surface-independent representations. Rather, longer training was associated with an increased surface-dependence effect, consistent with computational models that do not build surface-independent representations during training (e.g., Christiansen & Curtin, 1999; Dienes, Altmann, & Gao, 1999).

Finally, Study IV used opposition logic in order to pit controllable and automatic influences against each other in a generation task using grammars specifying first-order (Experiment 1) or second-order (Experiment 2) dependencies. Generation performance, as assessed by the degree to which generated sequences conformed to the grammars, was either low (Experiment 1) or at chance (Experiment 2). Although the results are somewhat inconclusive due to the low levels of generation performance, the knowledge participants were able to express in Experiment 1 was not applied untintentionally. In addition, Study IV illustrates in different ways how opposition logic can be implemented in AGL in the context of a generation task (similarly to the way in which these issues are often studied in the SRT task, Destrebecqz & Cleeremans, 2001; Fu et al., 2008).

Traditionally, implicit learning research has been rather divorced from other kinds of research, for example, memory and categorization research. However, some efforts have been made to relate implicit learning research to models, processes, and findings from other related paradigms (Higham, 1997b; Higham

& Brooks, 1997; Kinder & Shanks, 2001; Kinder et al., 2003; Pothos & Bailey, 2000; Vokey & Brooks, 1992; Whittlesea & Leboe, 2000). The potential unification of different paradigms associated with these efforts is promising, since progress in science is often a matter of unification of theory. The heuristic processing framework illuminated in Study I and II has the potential of providing a set of tools that may span over a wide array of research (see e.g., Whittlesea, 2003) and should therefore be investigated in more detail in the context of AGL and other implicit learning tasks.

One important issue concerns how different strategies and heuristics are coordinated in the course of knowledge application. Most computational models of AGL are more concerned with the knowledge acquisition part than the knowledge application part (e.g., Dienes et al., 1999; Kinder, 2000). Thus, it would be useful for future research on computational modelling to account more specifically for different effects resulting from differential strategic processing during knowledge application as well. This would enable a better understanding of the kind of computations that take place at test and how they are related to the knowledge that is expressed in implicit learning tasks. For example, the differential effects of analytic vs. non-analytic processing (Study I and II in this thesis; Kinder et al., 2003; Whittlesea et al., 1994) could be thought of partly as involving differential specificity with respect to activation of memory traces.[5]

A major source of controversy in AGL research has been whether the knowledge acquired is consciously available or not and how researchers should go about assessing the conscious status of knowledge in AGL (Dienes & Berry, 1997; Higham et al., 2000; Neal & Hesketh, 1997; Persaud et al., 2007; Perruchet & Vinter, 2002; A. S. Reber, 1989; Shanks & St. John, 1994; Tunney & Shanks, 2003a; Whittlesea & Dorken, 1997). Much research has been devoted to trying to dissociate different kinds of learning, in order to find evidence for qualitatively different learning and memory systems (e.g., Squire, 2004; Turner & Fischler, 1993).

One reason for the controversy surrounding the conscious status of knowledge is that the finding of unconscious knowledge (or dissociations between conscious and unconscious knowledge) has been taken by some researchers to provide support for a separate unconscious learning system. This debate is far from being settled and it naturally connects with deep underlying questions regarding both mental representation and what it means for knowledge to be conscious (Dienes & Perner, 1999; O´Brien & Opie, 1999; Perruchet & Vinter, 2002; Shanks & St. John, 1994). Arguments and data based purely on dissociation logic often turn out to be compatible with single-system models (Dunn, 2003; Kinder & Shanks, 2001, 2003; Shanks, 2005; Shanks & Perruchet, 2002; Whittlesea & Leboe, 2000; Whittlesea & Price, 2001) and the development of computational models will most likely continue to play an important role in understanding the modularity (or non-modularity) of learning, although the question of how to model awareness is not settled (for one suggestion, see Cleeremans, Timmermans, & Pasquali, 2007). Regardless of one´s stance with respect to these questions, the study of how different kinds of learning and testing conditions affect performance as assessed by different approaches to measure awareness is a

5 Exemplar models of memory and categorization seem apt in this context, because of their preservation of specific knowledge representations, leading to divergent effects during test probing (Hintzman, 1986). Although exemplar models were initially deemed problematic when applied to AGL (Dienes, 1992), more current research indicates that the viability of the models may depend on the coding scheme adopted (R. Jamieson, personal communication, January 11, 2008).

worthwhile research endeavour, especially if it can provide interesting ideas about the *function* of awareness (Cleeremans & Jiménez, 2002; Dienes & Perner, 1999; Perruchet & Vinter, 2002; Wegner, 2004). In this context, simply pointing to the existence of multiple learning and memory systems is of little help in explaining the diversity of performance and subjective attitudes that can be observed once various kinds of strategical and expectational factors are manipulated (Goldinger & Hansen, 2005; Kinder et al., 2003; Whittlesea, 2003, 2004).

Study I and II illustrate the importance and the potential of investigating how participants´ knowledge application and knowledge evaluation may be affected by different strategies for performing the task at hand. Specifically, Study I found that speeded responding, presumably encouraging non-analytic processing, was associated with an increased ability to distinguish between correct and incorrect decisions. Study II found that non-analytic processing instructions at test, as compared to analytic processing instructions, increased the participants´ ability to distinguish correct from incorrect decisions. It is certainly the case that one could find tasks and circumstances where analytic stategies are beneficial, so the conclusion is not that non-analytic processing is always beneficial in implicit learning. Rather, this may depend very much on the specific task at hand and on how the participants´ expectations interact with the specific strategy adopted. Future research could investigate how different kinds of expectations affect awareness in AGL in combination with different strategies. As emphasized by Whittlesea and Dorken (1997), it may be that awareness is itself a heuristic inferential activity. If so, then awareness could be expected to differ with respect to a variety of factors during the test situation. Clearly, varying such factors is essential to approaching an understanding of the different ways in which participants may develop different attitudes toward their own performance in implicit learning tasks.

In addition to the controversy regarding awareness, a great amount of controversy has surrounded the question of surface-independence in AGL research (Marcus et al., 1999; Perruchet & Vinter, 2002; A. S. Reber, 1989; Redington & Chater, 1996). The two controversies are connected though, because it is mainly abstraction without awareness that is controversial, for example, the notion of a powerful cognitive unconscious dedicated to abstraction toward surface-independent symbolic rules (cf. Perruchet & Vinter, 2002).

In contrast to the idea of automatic unconscious abstraction driving knowledge toward surface-independence (Manza & A. S. Reber, 1997) the results of Study III suggests that extended passive observational learning results in increasingly surface-dependent knowledge. Study III used a paradigm put forward by Conway and Christiansen (2006) and it would be interesting to investigate whether other kinds of learning conditions designed to promote active abstraction (e.g., through explicit instructions or integrative cues that might trigger abstractive processes) would result in more surface-independent knowledge over time. Such a result would be very much in line with the episodic-processing framework which suggests that the content of stored episodes is very much a function of attentional and intentional factors (Johnstone & Shanks, 2001; Whittlesea & Dorken, 1993; see also Pacton and Perruchet, 2008, for an attention-based framework). The episodic specificity of transfer of knowledge to different domains is an important research topic, since such transfer has traditionally been taken as a hallmark of cognitive abstraction. However, even highly symbolic tasks may show surprising degrees of perceptual dependence (Landy & Goldstone, 2007). Further development of computational models may elucidate the perceptual dependence of the knowledge supporting transfer in AGL. Particularly, it would be interesting to investigate which particular perceptual dimensions (if any) are dominant under different transfer conditions in AGL.

Study IV investigated the controllability of knowledge in AGL in the context of a generation task. Although the results were somewhat inconclusive because of low levels of generation performance, it should be emphasized that research on the SRT task has shown clear evidence of stable generation performance, even though the controllability of that performance is debated (Destrebecqz & Cleeremans, 2001; Fu et al., 2008; Wilkinson & Shanks, 2004). The training procedure used in Study IV obviously differs from the SRT task in many respects (e.g., memorization of exemplars rather than simply responding to individual locations) which may account for the more stable generation performance generally found in the SRT task. It would be interesting for future research to adopt different versions of the procedure and materials of Study IV (e.g., longer training, more constrained grammars, or a cued generation test) in order to investigate the effects on intentional vs. unintentional generation performance. Given that stable generation performance can be achieved under certain conditions, it would be interesting to further investigate whether such performance is affected by the same variables as generation performance in the SRT task. Such comparisons could offer important insights into common aspects associated with different implicit learning methods.

To sum up, the studies of this thesis provide insight into several aspects of AGL, including the heuristic nature of knowledge application and awareness, the surface-dependence associated with observational learning of regularities, and ways to investigate unintentional knowledge application in AGL. The studies also suggest potentially fruitful paths for future research, especially with respect to the episodic components of learning in combination with different heuristics. An important challenge in this regard will be to develop computational models able to implement restricted parts of the variability associated with different kinds of strategies in artificial grammar classification. This has the potential of unifying further findings from research on AGL, statistical learning, categorization, and memory research.

References

Allwood, C. M., Granhag, P. A., & Johansson, H. (2000). Realism in confidence judgments of performance based on implicit learning. *European Journal of Cognitive Psychology, 12,* 165-188.

Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 899–912.

Bayley, P. J., Frascino, J. C., & Squire, L. R. (2005). Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature*, *436*, 550-553.

Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2006). How to hit Scylla without avoiding Charybdis: Comment on Perruchet, Tyler, Galland, and Peereman (2004). *Journal of Experimental Psychology: General*, 135, 314-326.

Boucher, L., & Dienes, Z. (2003). Two ways of learning associations. *Cognitive Science*, *27*, 807-842.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*, 316–323.

Buchner, A. (1994). Indirect effects of synthetic grammar learning in an identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 550-566.

Chang, G. Y., & Knowlton, B. J. (2004). Visual feature learning in artificial grammar classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 714-722.

Channon, S., Shanks, D. R., Johnstone, T., Vakili, K., Chin, J., Sinclair, E. (2002). Is implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia*, *40*, 2185-2197.

Christiansen, M. H. & Curtin, S. L. (1999). The power of statistical learning: No need for algebraic rules. In *The Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 114-119). Mahwah, NJ: Lawrence Erlbaum.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406-416.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: an empirical, philosophical, and computational consensus in the making* (pp. 1-40). Hove, U.K.: Psychology Press.

Cleeremans, A., Timmermans, B., & Pasquali, A. (2007). Consciousness and metarepresentation: A computational sketch. *Neural Networks*, *20*, 1032-1039.

Cock, J. (2005). Completely ruled out? Using response time data to investigate biconditional artificial grammar learning. *European Journal of Cognitive Psychology*, *17*, 770-784.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905-912.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the Process Dissociation Procedure. *Psychonomic Bulletin & Review*, *8*, 343-350.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. In R. Banerjee & B. Chakrabarti (Eds.), *Progress in Brain Research, 168*, 49-64. Amsterdam, The Netherlands: Elsevier.

Dienes, Z. (1992). Connectionist and memory array models of artificial grammar learning. *Cognitive Science, 16, 41-79.*

Dienes, Z., & Altmann, G. T. M. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D.C. Berry (Ed.), *How implicit is implicit learning?* (pp. 107-123). Oxford, UK: Oxford University Press.

Dienes, Z., Altmann, G. T. M., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science, 23*, 53–82.

Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21, 1322-1338.*

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review*, *4*, 3-23.

Dienes, Z., & Fahey, R. (1995). The role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1322-1328.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735-808.

Dienes, Z., & Perner, J. (2004). Assumptions of a subjective measure of consciousness: Three mappings. In R. Gennaro (Ed.), *Higher order theories of consciousness* (pp 173-199). Amsterdam, The Netherlands: John Benjamins Publishers.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, *69*, 338-351.

Dunn, J. C. (2003). The elusive dissociation. *Cortex*, *39*, 177-179.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179-211.

Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, *105*, 577-614.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage, 32*, 956-967.

French, R. M., & Cleeremans, A. (Eds.). (2002). *Implicit learning and consciousness: an empirical, philosophical, and computational consensus in the making*. Hove, U.K.: Psychology Press.

Fu, Q., Fu, X., & Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness and Cognition*, *17*, 185-202.

Gaillard, V., Vandenberghe, M., Destrebecqz, A., & Cleeremans, A. (2006). First- and third-person approaches in implicit learning research. *Consciousness and Cognition*, *15*, 709-722.

Goldinger, S. D., & Hansen, W. A. (2005). Remembering by the seat of your pants. *Psychological Science*, *16*, 525-529.

Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, *33*, 154-207.

Gomez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote learning in language-learning infants. *Psychological Science*, *17*, 670-674.

Gomez, R. L., & Gerken, L. (1999). Artificial language learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109-135.

Gomez, R. L., Gerken, L. A., & Schvaneveldt, R. W. (2000). The basis of transfer in artificial grammar learning. *Memory and Cognition*, *28*, 253-263.

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33, 394-406.*

Helman, S., & Berry, D. C. (2003). Effects of divided attention and speeded responding on implicit and explicit retrieval of artificial grammar knowledge. *Memory & Cognition*, 31, 703-714.

Higham, P. A. (1997a). Chunks are not enough: The insufficiency of feature frequency-based explanations of artificial grammar learning. *Canadian Journal of Experimental Psychology*, *51*, 126–137.

Higham, P. A. (1997b). Dissociations of grammaticality and specific similarity effects in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 1029-1045.*

Higham, P. A., & Brooks, L. R. (1997). Learning the experimenter´s design: Tacit sensitivity to the structure of memory lists. *The Quarterly Journal of Experimental Psychology*, *50(A)*, 199-215.

Higham, P. A., Vokey, J. R., & Pritchard, J. L. (2000). Beyond dissociation logic: Evidence for controlled and automatic influences in artificial grammar learning. *Journal of Experimental Psychology: General, 129, 457-470.*

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.

Jimenez, L. (Ed.) (2003). *Attention and implicit learning*. Amsterdam, The Netherlands: John Benjamins Publisher.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118,* 126-135.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118,* 115-125.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–47). Mahwah, NJ: Erlbaum.

Jamieson, R. K., & Mewhort, D. J. K. (2005). The influence of grammatical, local, and organizational redundancy on implicit learning: An analysis using information theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 9-23.

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, *310*, 116-119.

Johnstone, T., & Shanks, D. R. (1999). Two mechanisms in implicit artificial grammar learning? Comment on Meulemans and Van der Linden (1997). *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 524-531.

Johnstone, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology*, *42*, 61-112.

Kelly, S. W., & Wilkin, K. (2006). A dual-process account of digit invariance learning. *The Quarterly Journal of Experimental Psychology*, *59*, 1664-1680.

Kinder, A. (2000). The knowledge acquired during artificial grammar learning: Testing the predictions of two connectionist models. *Psychological Research*, *63*, 95-105.

Kinder, A., & Assmann, A. (2000). Learning artificial grammars: No evidence for the acquisition of rules. *Memory and Cognition*, *28*, 1321-1332.

Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience*, *13*, 648-669.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review*, *110*, 728-744.

Kinder, A., Shanks, D. R., Cock, J. & Tunney, R. J. (2003). Recollection, fluency, and the explicit/implicit distinction in artificial grammar learning. *Journal of Experimental Psychology: General, 132,* 551-565.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 79-91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 169-181.

Kuhn, G., & Dienes, Z. (2005). Implicit learning of non-local musical rules: Implicitly learning more than chunks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1417-1432.

Kuhn, G., & Dienes, Z. (2006). Differences in the types of musical regularity learnt in incidental- and intentional-learning conditions. *Quarterly Journal of Experimental Psychology*, *59*, 1725-1744.

Kuhn, G., & Dienes, Z. (2008). Learning non-local dependencies. *Cognition*, *106*, 184-206.

Kurilla, B. P., & Westerman, D. L. (2008). Processing fluency affects subjective claims of recollection. *Memory and Cognition*, *36*, 82-92.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General*, *135*, 162-183.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 720-733.

Leboe, J. P., & Whittlesea, B. W. A. (2002). The inferential basis of familiarity and recall: Evidence for a common underlying process. *Journal of Memory and Language*, *46*, 804-829.

Lieberman, M. D., Chang, G. Y., Chiao J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *Journal of Cognitive Neuroscience,* 16, 427-438.

Litman, L., & Reber, A. S. (2002). Rules, abstractions, and evolution. *Behavioral and Brain Sciences*, *25*, 345-346.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149-157.

Lotz, A., & Kinder, A. (2006a). Classification and recognition in artificial grammar learning: Analysis of receiver operating characteristics. *The Quarterly Journal of Experimental Psychology*, *59*, 667-682.

Lotz, A., & Kinder, A. (2006b). Transfer in artificial grammar learning: The role of repetition information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 707-715.

Love, B. C., & Gureckis, T. M. (2007). Models in search of a brain. *Cognitive, Affective, and Behavioral Neuroscience*, *7*, 90-108.

Manza, L., & Reber, A. S. (1997). Representing artificial grammars: Transfer across stimulus forms and modalities. In D.C. Berry (Ed.), *How implicit is implicit learning?* (pp. 73–106). Oxford, England: Oxford University Press.

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT press.

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, *18*, 387-391.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–79.

McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences*, *3*, 166-168.

Merikle, E. M. (2003). Interpreting dissociations: The issue of task comparability. *Cortex*, *39*, 174-176.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1007–1028.

Neal, A., & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin & Review*, *4*, 24-37.

Newell, B. R., & Bright, J. E. H. (2001). The relationship between the structural exposure effect and the implicit learning process. *Quarterly Journal of Experimental Psychology, 54,* 1087 –1104.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Nissen, M. J., Bullemer, P. T. (1987). Attentional requirements for learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1-32.

Norman, E., Price, M. C., & Duff, S. C. (2006). Fringe consciousness in sequence learning: The influence of individual differences. *Consciousness and Cognition*, *15*, 723-760.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*, 247-255.

O´Brien, G., & Opie, J. (1999). A connectionist theory of phenomenal experience. *Behavioral and Brain Sciences*, *22*, 127-148.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and non-adjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 80-96.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130*, 401-426.

Pena, M., Bonatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604-607.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 223-228.

Perruchet, P. (2005). Statistical approaches to language acquisition and the self-organizing consciousness: A reversal of perspective. *Psychological Research*, *69*, 316-329.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: Two approaches, one phenomenon. *Trends in Cognitive Sciences*, *10*, 233-238.

Perruchet, P., & Peereman, R. (2004). The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics*, *17*, 97-119.

Perruchet, P., Peereman, R., & Tyler, M. D.

(2006). Do we need algebraic-like computations? A reply to Bonatti, Pena, Nespor, and Mehler (2006). *Journal of Experimental Psychology: General*, *135*, 322-326.

Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning non-adjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*, *133*, 573-583.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences*, *25*, 297-330.

Perruchet, P., Vinter, A., Pacteau, C., & Gallego, J. (2002). The formation of structurally relevant units in artificial grammar learning. *The Quarterly Journal of Experimental Psychology*, *55A*, 485-503.

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*, 257-261.

Pothos, E. M. (2005). Expectations about stimulus structure in implicit learning. *Memory and Cognition*, *33*, 171-181.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin, 133*, 227-244.

Pothos, E. M., & Bailey, T. M. (2000). The role of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 847-886.

Ratcliff, R., McKoon, G., & Van Zandt, T. (1995). Process dissociation, single-process theories, and recognition memory. *Journal of Experimental Psychology: General*, *124*, 352–374.

Reber, A. S. (1967). Implicit learning of artifical grammars. Journal *of Verbal Learning and Verbal Behavior*, *6*, 855-863.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, *81*, 115-119.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235.

Reber, A. S. (1997). How to differentiate implicit and explicit modes of acquisition. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 137-160). Hillsdale, NJ: Erlbaum.

Reber, R., Winkielman, P., & Schwarz, N. (1999). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*, 45-48.

Redington, M. (2000). Not evidence for separable controlled and automatic influences in artificial grammar learning: Comment on Higham, Vokey, and Pritchard (2000). *Journal of Experimental Psychology: General*, *129*, 471–475.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology:General*, *125*, 123-138.

Redington, M., & Chater, N. (2002). Knowledge representation and transfer in artificial grammar learning. In R. M. French & A. Cleeremans (Eds.), *Implicit Learning and Consciousness: An empirical, philosophical and computational consensus in the making* (pp. 121–143). Hove, UK: Psychology Press.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44:*563-575.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.

Rouder, J. N., & Ratcliff, R. (2006). Comparing exemplar- and rule-based theories of categorization. *Current Directions in Psychological Science*, *15*, 9-13.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*, 110-114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926-1928.

Sanna, L. J., & Schwarz, N. (2007). Metacognitive experiences and hindsight bias: It´s not just the thought (content) that counts! *Social Cognition*, *25*, 185-202.

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential basis of metamemory. *Current Directions in Psychological Science*, *6*, 132 - 137.

Seidenberg, M. S., & Elman, J. L. (1999). Networks are not "hidden rules." *Trends in Cognitive Sciences, 3*, 288-289.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 592-608.

Seth, A. K. (in press). Theories and measures of consciousness develop together. *Consciousness and Cognition.*

Shanks, D. R. (2005). Implicit learning. In K. Lamberts and R. Goldstone (Eds.), *Handbook of Cognition* (pp. 202-220). London: Sage.

Shanks, D. R., Johnstone, T., & Kinder, A. (2002). Modularity and artificial grammar learning. In R. M. French & A. Cleeremans (Eds.), *Implicit Learning and Consciousness*: *An empirical, philosophical and computational consensus in the making* (pp. 93-120). Hove, U.K.: Psychology Press.

Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin and Review*, *9*, 362–367.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367-447.

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

Stadler, M. A., & Frensch, P. A. (Eds.) (1998). *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage Publications.

Tunney, R. J. (2003). Implicit and explicit knowledge decay at different rates: A dissociation between priming and recognition in artificial grammar learning. *Experimental Psychology*, *50*, 124-130.

Tunney, R. J. (2005). Sources of confidence in implicit cognition. *Psychonomic Bulletin and Review*, *12*, 367-373.

Tunney, R. J., & Altmann, G. T. M. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 614–639.

Tunney, R. J., & Shanks, D. R. (2003a). Subjective measures of awareness and implicit cognition. *Memory & Cognition*, *31*, 1060-1071.

Tunney, R. J., & Shanks, D. R. (2003b). Does opposition logic provide evidence for conscious and unconscious processes in artificial grammar learning? *Consciousness and Cognition*, *12*, 201–218.

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552-564.

Turner, C. W., & Fischler, I. S. (1993). Speeded tests of implicit knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1165-1177.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328-344.

Vokey, J. R., & Higham, P. A. (2004). Opposition logic and neural network models in artificial grammar learning. *Consciousness and Cognition*, *13*, 565-578.

Vokey, J. R., & Higham, P. A. (2005). Abstract analogies and positive transfer in artificial grammar learning. *Canadian Journal of Experimental Psychology, 59*, 54-61.

Wegner, D. M. (2004). Precis of the illusion of conscious will. *Behavioral and Brain Sciences*, *27*, 649-692.

Werth, L., & Strack, F. (2003). An inferential approach to the knew-it-all-along phenomenon. *Memory*, *11*, 411-419.

Whittlesea, B. W. A. (2002). Two routes to remembering (and another to remembering not). *Journal of Experimental Psychology: General*, *131*, 325-348.

Whittlesea, B. W. A. (2003). On the construction of behavior and subjective experience: The production and evaluation of performance. In J. S. Bower & C. Marsolek (Eds.), *Rethinking implicit memory* (pp. 239-260). Oxford: Oxford University Press.

Whittlesea, B. W. A. (2004). The perception of integrality: Remembering through the validation of expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 891-908.

Whittlesea, B. W. A., Brooks, L. R., & Westcott, C. (1994). After the learning is over: Factors controlling the selective application of general and particular knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 259-274.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General*, *122*, 227-248.

Whittlesea, B. W. A., & Dorken, M. D. (1997). Implicit learning: Indirect, not unconscious. *Psychonomic Bulletin and Review*, *4*, 63-67.

Whittlesea, B. W. A., & Leboe, J. P. (2000). The heuristic basis of remembering and classification. *Journal of Experimental Psychology: General*, *129*, 84-106.

Whittlesea, B. W. A., & Price, J. (2001). Implicit/explicit memory versus analytic/nonanalytic processing: Rethinking the mere exposure effect. *Memory and Cognition, 29,* 234-246.

Whittlesea, B. W. A., & Wright, R. L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequenses. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 181-200.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 354-369.

Wright, R. L., & Whittlesea, B. W. A. (1998). Implicit learning of complex structures: Active adaptation and selective processing in acquisition and application. *Memory and Cognition*, *26*, 402–420.

Yates, J. F. (1994). Subjective probability accuracy analysis. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 381-410). Chichester, England: Wiley.

Zaki, S. R., & Nosofsky, R. M. (2007). A high-distortion enhancement effect in the prototype-learning paradigm: Dramatic effects of category learning during test. *Memory and Cognition*, *35*, 2088-2096.