The effect of overloading the memory capacity on rule learning

Thesis Research Master Linguistics: The Study of the Language Faculty – July 2017

Jessica van Schagen 3986829

Supervisors: Sergey Avrutin & Frank Wijnen

Abstract

The aim of this paper was to study the effect of overloading the brain's channel capacity, using a secondary memory task, on rule induction. The main hypothesis was that overloading the channel would lead to an increase in rule induction. A dual task paradigm was used to overload the channel: an artificial grammar task was used to study rule induction and a secondary task was used to overload the channel. Two groups of participants were compared: one group with a secondary memory task (the experimental group) and one group with a secondary attention distracter task (the control group). The results showed that the experimental group performed worse than the control group on all test items. This contradicted the hypothesis. A possible explanation could be that the memory task was more demanding than the attention distracter task, leading to less explicit learning of the artificial language in the experimental group compared to the control group (as participants were more distracted). Compared to previous studies using the same artificial language, the control group seemed to perform better on the rule induction task. This could mean attention underlies the brain's channel capacity. Independent measurements to account for individual differences in memory capacity and pattern recognition capacity were not significant as main effects. The results of this study can increase understanding of the brain's channel capacity and language learning in general.

Acknowledgements

I would like to thank my first supervisor Sergey Avrutin for his advice and class on Information Theory, in which I thought of my research questions. I would also like to thank my second supervisor Frank Wijnen for his continuous advice and support in the lab meetings. Furthermore, I would like to thank Silvia Radulescu for her support throughout the process; without her this thesis would not be there. I would furthermore like to thank the members of the babylab meetings for their feedback, Efi Giannopoulou for her feedback and support, especially during the test phase (thank you for the KitKats) and Eveline Sprangers for her feedback and assistance throughout the entire process.

Table of contents

List of figures

Introduction

People all over the world use language, whether auditory or visual, to communicate. In ordinary circumstances people start to learn these languages in the womb. What motivates humans to learn languages, especially at such a young age, has been up for debate for decades, if not centuries (Nowak, Komarova, & Niyogi, 2001). Is it an innate Universal Grammar given at birth, or do we for instance have special domain general mental capacities that help us tune in to, for example, the rules of language? This thesis will focus on mental capacities such as memory, attention and rule induction, and the effect of overloading some of these capacities on our ability to learn the rules of a language.

Theoretical framework

When learning a language, a speaker must remember not only the words of the language (the lexicon) but also the order in which to place these words (the grammar). The word order of languages is not random; there are rules. An example of such a rule in both English and Dutch, is that a main clause consisting of the categories noun, finite verb and adverb must be placed in that order: noun + finite verb + adverb. Examples of this particular word order can be seen in 1 and 2.

1. Johnny eats quickly.
2. Sarah laughs softly.

This study will examine what underlying mental capacities are used in finding the rules of a language, building on previous research by Radulescu, Wijnen and Avrutin (in prep) studying the effect of input complexity on rule induction using an artificial grammar task.

Radulescu et al. discussed two types of rule induction: item-bound generalisations and category-based generalisations. Item-bound generalisations are concerned with the relations between perceptual features of words. Radulescu et al. considered an example of this to be the perceptual item "-ed" that has to be added to any regular verb to transform them to the past tense form. Category-based generalisations deal with abstract variables (and not specific items). An example for this would be the noun, finite verb, adverb word order previously mentioned. Category-based generalisations do not deal with the specific nouns ("Johnny" or "Sarah"), finite verbs ("eats" or "laughs") or adverbs ("quickly" or "softly"), but with the abstract categories of NOUN, FINITE VERB or ADVERB. Such an abstract category can take many values, but the word order (the rule) stays the same. Radulescu et al. studied whether these item-bound generalisations (dealing with perceptual features of specific items) and category-based generalisations (dealing with abstract categories) were a product of the same underlying mechanism, and if so, what the factors were that triggered the use of one form of encoding over the other.

Information theory

Radulescu et al. addressed their questions using Shannon's Information Theory (1948), using the notions of channel capacity and entropy. According to Shannon, the channel is a "medium used to transmit the signal from a transmitter to receiver" (p. 7). A visual illustration of Shannon's model of a general communication system can be seen in Fig. 1. In Shannon's original work (1948), the channel was described as anything that can transmit a signal, such as a wire or radio frequencies. The signal would travel from an information source, via a transmitter, to a receiver and would end up at its destination. This journey would travel via the channel, which has a limited capacity. The noise source is used by Shannon, but will not be used for this paper as it is not manipulated in the experiment.

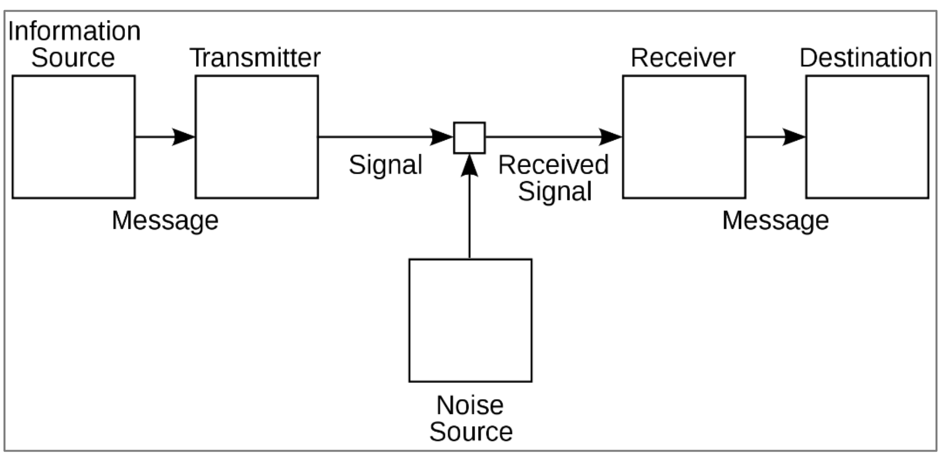

Figure 1. Shannon's model of a general communication system.

Shannon studied communication and information produced by a process. According to Shannon, the message sent in communication is one selected from a set of possible messages as the system, the speech organs and the brain for speech for example, must be built in such a way that they can produce each message desirable. To calculate how much information is produced in the process, the entropy formula from information theory can be used to calculate the complexity of a message. Entropy can be used to calculate the amount of uncertainty in a set (Goldsmith 2000;

2007). Entropy is concerned with the total number of items in the input, the number of individual items in the input and the probabilities of the individual items. The formula to calculate entropy is given below where H = entropy, Σ = the sum, x_i = a random variable x with n values and p = the probability.

$$H(x) = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

Shannon found out that the number of errors in a signal were not determined by the speed at which the message was sent, but by the complexity of the message (Van Ewijk & Avrutin, 2016). This entropy is expressed in bits and it is suggested that the channel capacity has a limited amount of bits it can process per second. If the message sent has more bits than the channel capacity allows for, errors can occur. An illustration on how to use the entropy formula to calculate the amount of bits in a set can be seen in Appendix A.

According to the formula, increasing the number of items in a set (such as the number of words in a language) increases the entropy of the language (if the probabilities of the items are equal), and changing the probability of items in a set from equal to unequal decreases the entropy.

Channel capacity

In Radulescu et al.'s paper, the channel described by Shannon is thought to be something physical, present in the human brain. The authors use this notion of a limited channel capacity to prove their hypothesis that the two forms of encoding (item-bound and category-based generalisations) are the product of the same underlying mechanism and that the complexity of the input (measured in entropy) triggers the change from one form of encoding to the other. They mention that item-bound generalisations are concerned with memorising specific items and their perceptual features, such as the phonotactic information of the word. An example would be the need to add the perceptual item "-ed" behind any regular verb to make a past tense form. Radulescu et al. mention that statistical learning, dealing with the co-occurrence probabilities and transitional probabilities, may be involved. However, as this learning mechanism is concerned with specific items, it cannot account for the generalisations to novel input (as the learner did not learn probabilities of those items). To do this, a learning mechanism must encode the novel input as a category to let the learner's mind know how to use it. Category-based generalisations, however, can generalise to novel input as they deal with variables. The previously used example of the noun-finite verb-adverb word order would be an example of this. Using examples from child studies Radulescu et al. reason that children must have two types of encoding: encoding statistical regularities (item-bound encoding) or encoding abstract rules (category-based encoding). They hypothesised that the two types of encoding (item-bound and category-based generalisations) are not independent, but the product of one underlying learning mechanism, and that the input complexity (measured in entropy) and the limited channel capacity trigger a gradual change from one type of encoding to the other. Should the channel capacity be too small to make item-bound generalisations, the learner is hypothesised to move to category-based generalisations to prevent overloading of her/his channel capacity. They further hypothesise that the channel and its capacity (measured in entropy) are the link between the two encoding systems, triggering the use of one encoding system over another if the input complexity exceeds the channel capacity. The channel could be seen as the neural connections in the brain, responsible for the sending of information, which can mature and eventually decline with age.

When talking about the channel capacity, both Shannon and Radulescu et al. did not discuss the number of channels in the brain. There is a possibility that there are more than one channel; one for each cognitive capacity for example. This uncertainty does pose problems theoretically, as the capacity per channel may differ and as participants may use different channels for different tasks.

Radulescu et al. study

To test their hypotheses, Radulescu et al. designed two experiments, examining the effect of input complexity on rule induction using an artificial grammar. They did this by changing the input complexity, whilst intending to keep the channel capacity the same (by using participants of similar ages and educational backgrounds). In their first experiment, they made a language with an XXY rule[1] consisting of Dutch-sounding (but meaningless) syllables, which they exposed participants to in three training phases. Participants were questioned about their knowledge of the language in small intermediate test phases after each training phase and a final test at the end. Radulescu et al. made three languages: a low entropy language, a medium entropy language and a high entropy language. Using the entropy formula given earlier, they calculated

[1] The XXY rule is a rule in which a sound (X) must be followed by the same sound (X) first and then followed by a different sound (Y). This a rule dealing with order (same-same-different sound), similar to the word order rule of noun-verb-adverb in language.

the entropy of each language and ended up with a total entropy of 3.5 bits for the low entropy language, 4 bits for the medium entropy group and 4.5 bits for the high entropy group[2]. Experiment 2 used a similar way of constructing three new languages, only now with more different syllables, increasing the total number of words and thus changing the entropy[3]. This created three new entropy condition of 2.8 bits, 4.26 bits and 4.8 bits. See Appendix B for the syllables used and the calculations of the entropy.

During the 3 intermediate test phases and the final test phase the participants were asked to judge whether a word could be correct in the language they had heard. There were four types of words they were asked to judge, each with a correct or incorrect response according to the XXY rule of the artificial language:

Test item	Correct/incorrect according to the XXY rule
Type 1) XXY_trained_syllables	Correct
Type 2) X_1X_2Y_trained_syllables[4]	Incorrect
Type 3) XXY_untrained_syllables	Correct
Type 4) X_1X_2Y_untrained_syllables	Incorrect

The X_1 and X_2 indicate that the X's were different from each other, but they were both X's (in case of type 2 they were X's throughout the input) and never used as Y's. Participants who would (perhaps not consciously) make item-bound generalisations were expected to accept type 1, but reject type 3 as it was unfamiliar. Participants making category-based generalisations were expected to recognise the XXY pattern and thus accept both type 1 and type 3 words. Type 2 was expected to be rejected by people making only item-bound generalisations (as they would remember the syllables and their sequences) and people making only category-based generalisations (as they would reject words not conforming to the XXY pattern). Participants

[2] Radulescu et al. not only used the probabilities of the syllables but also those of chunks of syllables, bigrams and trigrams, to calculate the total entropy as previous studies showed that knowledge about these items can shape grammaticality judgements in artificial grammar tasks (Perruchet and Pacteau, 1990; Knowlton and Squire, 1994; Pothos, 2010).

[3] The low entropy condition of experiment 2 was slightly different in that it did not match each XX pair to a different Y-syllable, but to the same Y-syllables all the time. This gave 7 different XXY words that were repeated 4 times each. This was done to ensure a lower entropy condition (if each XX pair was matched to a different Y syllable it would give an entropy of 3.7 bits).

[4] In experiment 1, type 2 and type 4 were referred to as XYZ structures instead of X_1X_2Y structures. In the XYZ words, X and Y syllables were randomly assigned to either the X, Y or Z position. The X_1X_2Y structures was thought of to see if the responses in experiment 2 were influenced by the fact that X's in the input could switch to the Y-position in the XYZ structure. As the results of experiment 2 confirmed the results of experiment 1, the term X_1X_2Y will be used throughout this paper.

switching between both types of generalisations were expected to perform worse, as they neither fully remembered the syllables and their sequences, nor would they fully reject words not following the XXY pattern. They would thus accept type 2 words more often. Type 4 was expected to be rejected by all, as neither the syllables, the syllables and their sequences nor the pattern would look familiar to participants using any type of generalisation.

The results from both experiments confirmed the predictions. Participants in all entropy groups scored above 93% in accepting type 1 words. There seemed to be a linear correlation between entropy and acceptance of type 3: the higher the entropy, the higher the acceptance of type 3 words. Participants in the higher and lower entropy groups were more likely to reject type 2 words than those in the medium entropy groups; a higher input entropy prevented strong memory traces for participants in the medium entropy condition, while the entropy was not high enough to lead to fully formed category-based generalisations. These results indicate that a higher input complexity (expressed in entropy) leads to more category-based generalisations. In their discussion, Radulescu et al. propose to study the underlying cognitive capacities of the brain's channel capacity. They hypothesise that the channel is captured by working memory capacity (WMC) and pattern recognition capacity (PRC), as the WMC is concerned with the memorisation of items (item-bound generalisations) while the PRC is concerned with patterns of the language, such as word order.

Artificial grammars

Artificial grammar learning is a way to test language learning abilities without the risk of a participant having heard the language before. Artificial languages often study rules, or patterns, of a language without meaning. These patterns can be simple such as an XXY pattern, more complex such as non-adjacent dependencies, XYZ[5], or very complex such as the Markovian systems used by Reber (1989). They can be used to study the difference between explicit and implicit learning (the learning mechanism), or the contrast between for example rules and associations (the learning product, or knowledge) (Pothos, 2007). Learning is seen as a modification in the stimulus-response relation that emerges through repetition, concerned with the immediate task only, and that occurs as a consequence of an interaction with the

[5] In non-adjacent dependencies there is a relationship between two kinds of sounds, separated by a completely different and changing sound in between. For example, "mie-gaa-roo" and "mie-poo-roo". The "mie" and "roo" are always together, separated by a different sound (such as "gaa" or "poo"). If a participant was to judge the word "mie-gaa-lie" it should be judged as ungrammatical.

environment, based on the definitions of Lachman (1997) and Fiol and Lyles (1985). Thus, learning in an artificial grammar learning task is caused by an interaction with the environment (the training items) of the task and the stimuli in that environment and can cause a change in the relation with the formerly unknown stimuli and a required response (the test items). As Radulescu et al. (in prep) mentioned, probabilities in the input are concerned with statistical learning, and this was deemed not sufficient as a learning mechanism on its own as it cannot generalise to new items. The fact that item memorization takes place is proven by results showing that items from the training phase (or very similar to them) are more likely to be judged as grammatically correct (Redington & Chater, 1996). If this occurs even when rule abstraction takes place, it confirms the notion from Radulescu et al. that there is a gradual transition from item memorisation (or item-bound generalisations) and rule abstraction (or category-based generalisations), in which participants can use parts of both learning mechanisms at the same time. This shows that several test items must be used to study both the use of item-based generalisations and category-based generalisations in one participant.

According to Redington and Chater there are three types of abstraction. The first type is a trivial type, the kind of abstraction participants make in realising the word is a word, and not just some random sounds. The second type is an abstraction over surface properties (or perceptual characteristics). This would relate to people remembering only the structure of the items they heard. For the types used in Radulescu et al. this would mean accepting the XXY_trained_syllables words (type 1), but not the XXY_untrained_syllables words (type 3). The third type of abstraction "abstracts away from the specific vocabulary used in the training set" (p. 124), meaning that it does not deal with the perceptual features from the training items. This would be the category-based generalisations in Radulescu et al. predicted to occur with the acceptance of type 3 test items. The fact that Redington and Chater mention that acceptance of type 1 items is still some sort of abstraction (and not just memorisation), shows that in order to make any claims on the form of encoding used, a comparison between types must be made. For example: type 1 acceptance rates would not be enough to make a claim about item-bound generalisations if type 3 acceptance is not taken into consideration, as type 1 items could be judged as correct for both people making item-bound generalisations as people making category-based generalisations. In order to make statements about category-based generalisations the difference between type 1 and type 3 acceptance must thus be studied. Similarly, in order to make claims about the memorisation of syllables, the judgements of type 2 items need to be compared to those of type 1, as type 1 items are both concerned with memorisation of syllables and rule-learning.

Reber mentions that his Markovian systems are too complex to be learned “in one afternoon in a laboratory” (1989, p. 220) and that this was needed to study implicit learning. Should the system be easy enough to be learned consciously, the learning that would be studied is would not be implicit and explicit learning can take place[6]. Thus, the difficulty of the artificial language must be taken into consideration as it may change the type of learning that takes place. As the artificial grammar from Radulescu et al. (in prep) is simpler than the Markovian structure studied by Reber, and consciously learnable in one afternoon in a laboratory, this means that explicit learning is expected to take place. This could mean that the learning studied using such a language would consist of some form of explicit learning. According to Reber (1977) implicit learning takes place naturally, without conscious operations such as hypothesis testing. This seems to contradict studies that show that implicit learning takes place both with undivided attention and divided attention, but when looking closely, tasks such as a dual task still require the participant to pay attention to the environment (though perhaps not consciously to the training items). Studies indeed show that implicit learning still takes place in dual task experiments (in which the participant’s attention is divided between two tasks), though perhaps to a lesser extent than in a single task experiment (Cleeremans, Destrebecqz & Boyer, 1998). López-Barroso, Cucurell & Rodrígues-Fornells (2016) even measured the incidental learning (using an online implicit measurement) separately from explicit learning (using an offline measurement) and found that implicit (or incidental) learning indeed took place regardless of the amount of attention paid to the stimuli, while explicit learning depended on the amount of attention paid during the training phase. These results suggest that implicit and explicit learning could take place simultaneously in Radulescu et al.; implicit learning seems to occur naturally, regardless of the amount of attention paid to the stimuli in the training phase, while explicit learning only seems to take place when the stimuli are easy (defined by Reber as learnable in one afternoon in a laboratory). The addition of a secondary task may switch the learning in the rule induction task in Radulescu et al. to more implicit learning. The difficulty level of the secondary

[6] Implicit learning is expected to take place subconsciously, without conscious effort and to anyone who pays attention to stimuli. Explicit learning is more conscious, giving the participant more explicit knowledge about the language. If the participant is able to express what she/he learned from the language, this knowledge is explicit.

task must thus be taken into consideration if two groups are compared to make sure the form of learning (implicit or explicit) is similar in both groups. Should the form of learning not be the same, the participant in different groups could have different amounts of knowledge of the language and this could influence the results.

Memory

As working memory capacity (WMC) was assumed by Radulescu et al. (in prep) to be one of the factors of the brain's channel capacity, a clear definition of WMC must first be established. However, this is not easy. Memory is a concept with many definitions, depending on which theory is taken into consideration, and many models. To unify some of these theories to be able to refer to definitions in this paper, I will work with and combine three existing memory models discussing working memory and its relations with other memory parts: the Functional Framework of Baars and Gage (2007), the Model of the Working Memory from Baddeley (2000) and the Information-processing Approach to Memory and the Components of Long Term Memory from Baddeley, Eysenck and Anderson (2015). The model unifying these three models can be seen in Fig. 2. (p. 15). This model will now be explained using the example of artificial grammar learning. The input would be the words heard in a training phase. This input would continue to the sensory memory, more specifically the echoic (or auditory) memory. The model would then continue to the working memory. According to Baddeley et al. this can be seen as a "mental workspace" (p. 13), used to temporally keep things in mind and perform manipulations on information. The component of the model concerned with temporally keeping things in mind is referred to as Short Term Memory (STM). This short remembering is split into (at least) two segments according to Baars and Gage and Baddeley et al.: a phonological loop and a visuo-spatial sketchpad. For the purpose of the artificial grammar learning example, only the phonological loop is involved (the visuo-spatial sketchpad does approximately the same for visual and spatial information). The idea of the phonological loop is to repeat what you have heard in order to remember. The STM is the part of the working memory that Radulescu et al. referred to when they spoke of the involvement of the WMC on the channel capacity; passive, temporary storage. The manipulation of information does not play into pure memorisation, but deals with processes or operations on information such as, for example, adding or subtracting numbers. Lastly, the information from the input goes to the Long Term Memory (LTM), in which it can be stored under for example linguistic or semantic memory. The information about interests that are held in the LTM can direct someone's attention to items in the input. E.g. someone who likes animals may focus more on the dog than the wallpaper when visiting a neighbour. As the channel was originally seen by Shannon as something that transmits a signal, the entire track from the phonological input to the LTM can be seen as the channel. Keeping this in mind, it makes sense that the brain's channel capacity can be captured by the WMC (and specifically the STM).

Previous research shows indeed that WMC is related to cognitive processes such as rule-induction (Kyllonen & Christal, 1990; Willis, Barrasin, & McLaren, 2011; Baars & Gage, 2007). Willis et al. found that participants with a large WMC (as measured by the operation span task, the OSPAN) were able to find underlying rules (an opposites rule, in which the presence of two elements lead to an opposite reaction compared to the presence of one element), while participants with a small WMC could only find surface rules (an additive rule, in which the presence of two elements lead to the same reaction as the presence of one element). However, the study does not find how well participants with a large WMC are at finding rules in general, and this is needed to make generalisations about artificial grammar learning as those studies focus learning one rule (and not underlying rules versus surface rules). A second reason not to generalise the results of Willis et al. to rule induction in general, is that they used the OSPAN test, which involves a secondary processing component (De Jong, 2010). The OSPAN task would thus measure the entire WMC, and not just the STM component that is predicted to underlie item-bound generalisations in the discussion by Radulescu et al. A more suitable task to measure the STM capacity would be to use a digit span task (Baddeley et al., 2015). In such a task, participants would hear or see a sequence of digits and they would be asked to repeat these in the order in which they were given (a forward digit span task) or in the reversed order (a backward digit span task). The forward digit span task would measure the STM component, as participants are only asked to listen, possibly repeat the numbers in their mind, and repeat what is given; no manipulations or processes are present. The backward digit span task, in which the participants are asked to reverse the order of the digits they memorised, deals with the working memory, e.g. storing and manipulation (in this case reversing) of information.

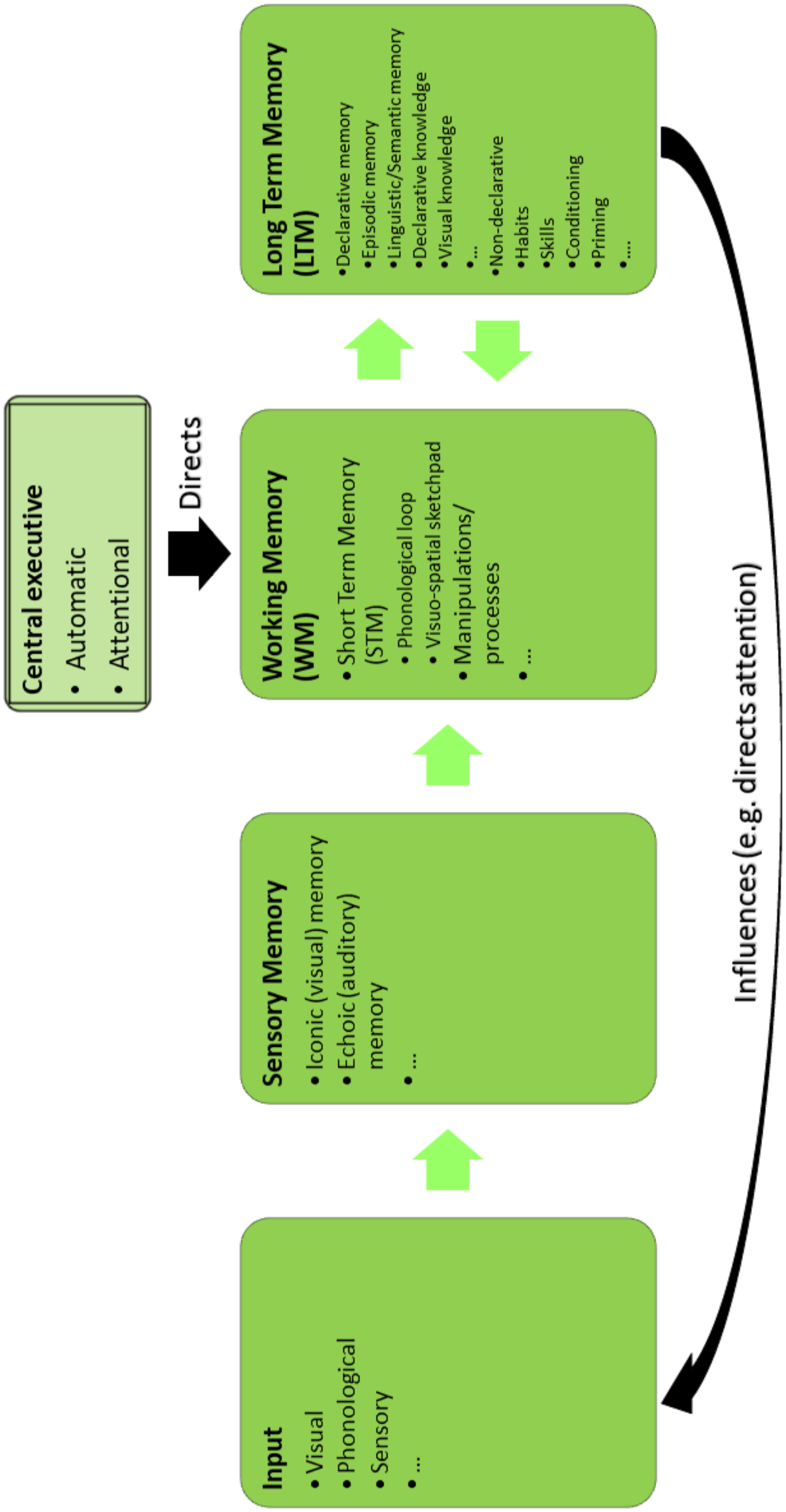

Figure 2. Unified memory model based on the memory models from Baars and Gage, Baddeley and Baddeley et al.

Pattern recognition capacity

Pattern Recognition Capacity (PRC) will be defined in this paper as the ability to find patterns (rules in languages), or the ability to make category-based generalisations (e.g. accepting XXY_new_syllables items words in Radulescu et al.). There are not many PRC tests. One well known test is the Raven's Standard Progressive Matrices, in which the participant has to find the missing piece of a pattern. The matrices in the Raven's Standard Progressive Matrices have different difficulty levels according to, for example, the dimensionality of the analogy (Anderson, 1990). According to Little, Lewandowsky and Griffiths (2012), these matrices test rule induction based on posterior probability. The fact that the Raven's Standard Progressive Matrices are of a different domain (vision) than language (auditory) should be of no concern as Altmann, Dienes and Goode (1995) found that the ability to abstract general rules could transfer to other domains (from auditory to vision). Taking the results from Altmann et al. into consideration, the Raven's Standard Progressive Matrices should be a good PRC test to test capacities used during an artificial grammar experiment.

Dual tasks

The point of a dual task paradigm is to divide participants' attention over two tasks, causing the participant to share her/his resources (or channel capacity) over two tasks, thus having less capacity for each individual task (Pashler, 1994). This links closely to the implicit versus explicit learning debate mentioned in the section Artificial Grammars. As the artificial language used by Radulescu et al. is simple, it is expected that participants learn the language more explicitly. The addition of a secondary task may lead the participants to direct their attention, and with it their capacity, to the secondary task. This could mean the artificial grammar might be learned more implicitly. To study whether the secondary task of the dual task paradigm has such an effect on rule induction, the results of this study will need to be directly compared to the study of Radulescu et al. For some studies a dual task paradigm can be problematic. This is because a system can be overloaded, especially if the same modalities are used (Pashler, 2016). However, overloading a system is the intention of this study, and therefore a dual task paradigm in the same modality (auditory) should be chosen.

The present study

This paper aims to study the unanswered questions posed in the discussion of Radulescu et al. (in prep) concerning the underlying cognitive capacities of the brain's channel capacity, in order to give more insight into the relatively unstudied brain channel capacity used for learning languages. Instead of increasing the input complexity with the aim of overloading the channel capacity, I will overload the channel capacity using a dual task paradigm to see if this method leads to a similar increase in category-based generalisations. This will be done by looking both at the working memory capacity and pattern recognition capacity. I will use one of the languages from Radulescu et al. My research questions are stated as follows:

> **Research question 1:**
> Will overloading the brain's channel capacity using a secondary memory task, to overload the memory capacity, lead to an increase in category-based generalisations?
>
> **Research question 2:**
> Will overloading the brain's channel capacity using a secondary pattern recognition task, to overload the pattern recognition capacity, lead to an increase in category-based generalisations?

To account for individual differences, participants' individual WMC and PRC will also be tested and taken into account. This is done in order to be able to better generalise results to the population. This leads to the following research questions:

> **Research question 3:**
> Does a person's independent working memory capacity (WMC) influence her/his tendency to make category-based generalisations?
>
> **Research question 4:**
> Does a person's independent pattern recognition capacity (PRC) influence her/his tendency to make category-based generalisations?

Hypotheses

To answer the research questions, theories about the channel capacity must be taken into consideration. If there is only one channel in the brain, the following hypotheses can be posed:

> **Hypothesis 1.1:**
> Adding a memory secondary task will overload the channel faster compared to a control secondary task, leading to an increase in category-based generalisations due to an overloaded channel capacity.

Hypothesis 2.1:
Adding a pattern recognition secondary task will overload the channel faster compared to a control secondary task, leading to an increase in category-based generalisations due to an overloaded channel capacity.

Should there be more than one channel, where the memory channel is independent from rule induction channel but pattern recognition is not (creating at least two channels then), the following hypotheses can be posed:

Hypothesis 1.2:
Adding a memory secondary task will overload the memory channel, but the rule induction channel will remain unchanged. Compared to the control group of Radulescu et al.'s experiment without a secondary task, the amount of category-based generalisations will be similar.

Hypothesis 2.2:
Adding a pattern recognition secondary task will overload the channel concerned with rule induction and pattern recognition, leading to an increase in category-based generalisations due to an overloaded channel capacity.

Should pattern recognition have no relation with rule induction and have an independent channel, similar to memory (thus creating at least three channels), the following hypotheses can be posed:

Hypothesis 1.3:
Adding a memory secondary task will overload the memory channel, but the rule induction channel will remain unchanged. Compared to the control group of Radulescu et al.'s experiment without a secondary task, the amount of category-based generalisations will be similar.

Hypothesis 2.3:
Adding a pattern recognition secondary task will overload the pattern recognition channel, but the rule induction channel will remain unchanged. Compared to the control group of Radulescu et al.'s experiment without a secondary task, the amount of category-based generalisations will be similar.

Should memory and pattern recognition capacities be unrelated to the channel capacity, the amount of attention or awareness to the stimuli might affect the participant's learning mechanism in terms of explicit learning versus implicit learning. The following hypotheses can be posed concerning explicit and implicit learning:

> **Hypothesis 1.4:**
> Adding a memory secondary task will cause the participant to pay less attention to the artificial language. She/he will therefore learn this language more implicitly and less explicitly. She/he will thus make less category-based generalisations. The more distracting the secondary task will be, the less attention the participant will pay to the language, the less category-based generalisations will be made.

> **Hypothesis 2.4:**
> Adding a pattern recognition secondary task will cause the participant to pay less attention to the artificial language. She/he will therefore learn this language more implicitly and less explicitly. She/he will thus make less category-based generalisations. The more distracting the secondary task will be, the less attention the participant will pay to the language, the less category-based generalisations will be made.

The following hypotheses will be posed about the independent memory and pattern recognition capacity:

> **Hypothesis 3:**
> A higher independent memory capacity will have the same effect of a bigger initial capacity of the channel with which memory is concerned. If memory is concerned with the same channel as rule induction, participants with a higher independent memory score will memorise more individual items and thus make less category-based generalisations.

> **Hypothesis 4:**
> A higher independent pattern recognition capacity will cause participants to look for patterns (or rules when language is concerned) faster. This will lead to an increase in category-based generalisations.

The results of this study will be able to shed some light on the brain's channel capacity, a relatively new field of research in linguistics. Taking into account attention, and independent capacity scores, the results of this study can shed more light on theories on memory and may assist in language learning applications. Due to time constraints only the secondary memory task will be studied; the pattern recognition secondary task should be studied in future research. This means hypotheses 2.1-2.4 cannot be confirmed or refuted in this paper.

Methods

Participants

57 Dutch-speaking adults were included (12 male and 45 female). The age range was 18-68[7] (M = 24, SD = 8.2). 14 additional people were tested but were excluded for technical problems (n = 4), having previous knowledge of what the test was about (n = 2), incorrect reading of instructions (n = 2) and hearing a longer warm up[8] (n = 6). Only healthy participants, with no known hearing problems or dyslexia were included. The study was approved by the ethics review board of the Uil-OTS. Participants were paid 10 euros for their participation.

Tasks

Rule induction task

Participants' ability to induce rules was measured using a dual task paradigm. This task consisted of two simultaneous parts: the main rule induction task and a secondary task to overload the channel capacity.

Main task: Rule induction task

The rule induction task used the artificial grammar language from Radulescu et al. with a total entropy of 2.8 bits. The choice was made to use the lowest entropy language. This was the condition in which Radulescu et al. thought the participants used their memory more than their pattern recognition abilities. If a higher entropy was chosen, in which participants already made category-based generalisations, there would be less room for the participants to increase in their use of category-based generalisations.

There were three training phases, three intermediate test phases after each training phase and one final test phase in the end. The intermediate tests were used to see if participants learned

[7] This large age range was thought to lead to more variation in independent memorisation and pattern recognition scores. More variation could mean the effect of the independent measurements could be studied better. Therefore, no participant was excluded on the basis of age.
[8] A pilot test was performed to examine the minimum length needed for the warm up.

more about the language as they were exposed to more words of the language. There were four test items types in the tests:

Test item	**Correct/incorrect according to the XXY rule**
Type 1) XXY_trained_syllables	Correct
Type 2) X1X2Y_trained_syllables[9]	Incorrect
Type 3) XXY_untrained_syllables	Correct
Type 4) X1X2Y_untrained_syllables	Incorrect

Type 1 items serves as a control, to make sure the participants paid attention to the rule induction task in this dual task paradigm and accept items that were heard before. Type 2 items are used to distinguish memorisation of syllables only (type 2 items) from memorisation of syllables and the pattern (type 1 items). Type 3 items are the main focus of this study as they test for category-based generalisations. A comparison between type 1 and type 3 items will provide information on item-based generalisations, as participants making those generalisations may accept XXY_trained_syllables words, but reject of XXY_untrained_syllables words. Type 4 items are a negative control, to see if participants also reject incorrect items. A comparison between type 2 and type 4 items may shed light on memory traces. Participants with memory traces of the syllables may incorrectly accept more type 2 items than type 4 items due to the known syllables.

In each training phase the participants heard 28 tokens from the artificial language consisting of 7 individual types, each repeated four times, all following the XXY pattern. In each intermediate test phase the participants were asked to judge if the word they heard would be correct in the language they had heard. They were asked to judge four words in each intermediate test phase, one of each type, and eight words in the final test, two of each type. For the grammatical judgement task a button box was used. All the words in the training and test phases were presented randomly.

Secondary task: Memory/Attention task

The secondary task consisted of listening to numbers and bleeps that participants would hear during the training phase. They heard the artificial language and the numbers and bleeps simultaneously. There were two secondary tasks: one for the control group (overloading

[9] In experiment 1, type 2 and type 4 were referred to as XYZ structures instead of X1X2Y structures. In the XYZ words, X and Y syllables were randomly assigned to either the X, Y or Z position. The X1X2Y structures was thought of to see if the responses in experiment 2 were influenced by the fact that X's in the input could switch to the Y-position in the XYZ structure. As the results of experiment 2 confirmed the results of experiment 1, the term X1X2Y will be used throughout this paper.

attention) and one for the memory group (overloading attention and memory). Both groups would eventually be compared to Radulescu et al.'s 2.8 bits group. However, as they did not use a dual task paradigm, comparing the memory overload group to the Radulescu et al. 2.8 bits group would not be a fair comparison as any effect that may be found could be caused by the memory overload or the addition of a secondary task in general. Thus, the attention group serves as a control for the memory group.

Participants were pseudo-randomly divided over two groups (N = 28 for the memory group, N = 29 for the control group). The memory task consisted of listening to two to four numbers followed by a “bleep” sound (described as “pieptoon” in Dutch). The participants were instructed to write down the one-but-last number they heard before the “bleep” sound on an answer sheet[10]. The last digit heard was not used as this may reflect the echoic memory more than the phonological loop in the STM. Because participants did not know how many numbers they would hear before the bleep, the task occupied their Short Term Memory (as they did not perform any manipulations on the numbers). An example of what the participants heard is given in 3.

3. 5 – **8** -4 – BLEEP – **1** – 0 – BLEEP – 7 – 3 – **5** – 1 - BLEEP

In this example the memory group had to write down “8 1 5” (presented in bold in 3).

The attention task consisted of listening to a string of two to four bleeps followed by a number. Participants were instructed to write down every number they heard on an answer sheet. They did not have to remember a string of numbers (as the memory group did). An example of what the participants heard is given in 4.

4. BLEEP – BLEEP – BLEEP – **6** – BLEEP – BLEEP – BLEEP – BLEEP – **1-** BLEEP – BLEEP – **3**

In this example the attention group had to write down “6 1 3” (presented in bold in 4).

---

[10] The number task was inspired by McAllister et al. (2001), in whose experiment participants had to name the digit they heard or do an N-back task, and Kantowitz and Knight (1976), in whose experiment people had to name the digit they heard as a secondary task in a dual task. The choice for digits as a secondary task (as opposed to tones or letters) was made as the digit span task as a dual task was used before by McDowell, Whyte and D'Esposito (1997) and a digit naming task was used before by Mcallister et al. and Kantowitz and Knight. A second reason not to use tones would be that Dutch is not a tone language and all the participants were Dutch native speakers.

The sounds from both tasks were presented through speakers. See Appendix C for the stimuli and test items used during the rule induction and secondary tasks.

A risk of dual task experiments is the bottleneck problem of dual tasking (Lien, Ruthruff, & Johnston, 2006). This problem is concerned with participants doing tasks alternatingly (thus leading to task switching), instead of simultaneously. To account for this problem both the scores of the primary task and the secondary task need to be taken into consideration to see if the participant did both tasks, and did not focus on the primary task only.

A female voice was used for both the artificial grammar items as the numbers used in the secondary task. This was done as female voices are generally more noticeable than male voices due to their higher pitch level, wider pitch range and possibly a greater vocal jitter. This was previously found to influence memory as participants paid more attention to the perceptual attributes of words than the content when words were spoken by female voice compared to male voices (Yang, Yang & Park, 2013). The artificial grammar task items and the secondary task items were spoken by different women to prevent the two sound streams from blurring into one sound stream.

Digit span task

To test their explicit STM, participants were exposed to a sequence of digits in a forward digits span task and they were asked to repeat them in the order that was given. They were told it was a memorisation task from the start. The participants listened to a female voice saying the digits. During the speaking phase the computer screen was empty. This was done to prevent participants from making mnemonic devices with digits on the screen. When the voice was done speaking the participants were asked to select which digits the voice said using their mouse to select digits that were presented on the screen. This continued until the participants made 2 mistakes in a sequence, after which the task stopped. The sound was presented through speakers.

Incidental memorisation task

To test their implicit memory, participants took part in an incidental memorisation task based on Stark and Okado (2013). During the incidental memorisation task, participants were at first not made aware that their memory would be tested according to the definition of incidental memorisation from Baddeley et al. (2015). Participants were first exposed to 30 nonsense words and asked to indicate whether the words referred to flowers, animals or tools based on what the words sounded like. After this categorisation phase, the participants were told their memory would be tested. They were asked to say whether they had heard the words presented to them in the memory phase before. Participants were then presented with 30 words, 15 of which they had heard in the categorisation phase and 15 of which they had never heard. The 15 words they had heard were randomly taken from the categorisation phase. All participants heard the same words. They were asked to indicate whether they had heard the words before by pressing “yes” or “no” on the screen using the mouse. See Appendix D for stimuli used in the incidental memorisation task. All items in the categorisation and memorisation phase were presented randomly. The sound was presented through speakers throughout the test.

Raven's standard progressive matrices

To test their pattern recognition ability, participants were asked to complete the Raven’s Standard Progressive Matrices. This test was done on paper. Participants were given a booklet with a pattern on each page. For each pattern, one part was missing and the participants could choose between 6-8 answers which pattern they thought would complete the incomplete pattern. They could write down their number on an answer sheet. The Raven’s Standard Progressive Matrices consist of 5 parts, A-E, increasing in difficulty. Each part consists of 12 questions, of which each first question is easy with the level of ease decreasing as the questions progress. The participants were given 35 minutes to answer all 60 questions. They were informed of the time at 20 and 30 minutes. If the participant completed the task early, they could let the experimenter know. An example of a Raven’s Standard Progressive Matrix can be seen in Appendix E.

Experiment session

All tasks of the experiment took place in a soundproof booth. Participants were welcomed in the lab by a female experimenter and given an information letter to read. When they agreed they had to sign an informed consent form. They were informed they would do four tasks.

The first task was the rule induction task combined with the secondary task. This was done so no other task could influence their results and make them more focussed on the rules (patterns) or individual items (using memory). Participants were informed they had to do two tasks at the same time. They were first given instructions on paper to read and could ask questions if they wanted to. Participants were asked not to cheat for either the rule induction task or the number task by taking notes. They were then given a warm up of 30 seconds of the secondary task only, to make sure they understood what to do. After the warm up there was time to ask questions again. When everything was understood, the dual task set-up would start. After the dual task, the experimenter would walk in and take away the instructions, the answer sheet of the number task and the pen. This was done so participants would not be able to make notes, and thus cheat, during the second and third task.

The second task was the forward digit span task. Participants were informed that the instructions would appear on the screen. When they understood everything they could click on the screen to continue and start the task.

The third task was the incidental memorisation task. The choice was made to perform this task after the digit span task (and not the rule induction task), to prevent the participants from trying to learn the language (as that would test explicit learning and not implicit or incidental learning). The participants were given instructions on paper. When everything was clear they could start the experiment by pressing any button on the button box.

The fourth task was the Raven's Progressive Matrice task. This were chosen as the last task as the participants could complete the test earlier. The participants were first given instructions on paper. If everything was clear the participant was told that the experiment leader would come in after 20 minutes and 30 minutes to inform the participant of the time. The participant was then free to start and had 35 minutes to complete the matrices. If the participant was done earlier, she/he could let the experiment leader know.

After the Raven's Progressive Matrices was completed the experiment leader asked the participants a few questions for feedback on what the participant noticed about the language and how they felt throughout. The participant was then asked to sign a sign-up sheet and she/he received 10 euros for participation. The experiment sessions lasted a maximum of 60 minutes.

Results

The performance in the secondary task was examined to see if participants paid attention. The percentage correct answers ranged between 76-100% (M = 98.2, SD = 4.7). No participant was excluded on the basis of not paying attention as they all performed above chance (50%). The range for the digit span was 4-10 (M = 6.72, SD = 1.37). The incidental memorisation scores were calculated using Signal Detection Theory (Macmillan & Creelman, 2004). The d' assesses how well the noise (incorrect rejections and incorrect acceptances) and the signal (correct acceptances and correct rejections) are discriminated. The d' range for the incidental memorisation task was 0.34-2.61 (M = 1.67, SD = 0.60). The Raven's Standard Progressive

Matrices test scores were calculated using percentiles for age groups, based on the percentile table for adults in the USA (*Raven*, 2006, p. 95). The range for the Raven's test percentiles was 0-100 (M = 74, SD = 21.8).

To assess the performance on the rule induction task, type 1 and 3 items were scored as correct if they were accepted and type 2 and 4 items were scored as correct if they were rejected. The accuracy scores for each type are presented in Fig. 3 and Table 1 and the accuracy scores per test are presented in Fig. 4. and Table 2.

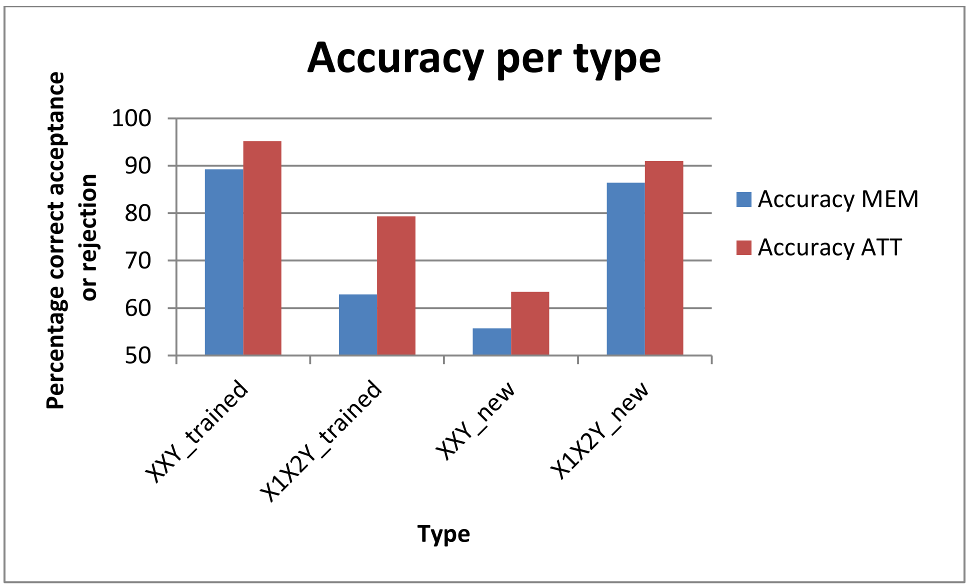

Figure 3. Performance per type (for all tests) for the experimental condition (MEM) in blue and the control condition (ATT) in red. Type 1 (XXY_trained) and type 3 (XXY_untrained) were scored as correct if they were accepted and type 2 (X1X2Y_trained) and type 4 (X1X2Y_untrained) were scored as correct if they were rejected.

	Memory (MEM)		Attention (ATT)	
	Mean(%)	**SD**	**Mean(%)**	**SD**
XXY_trained_syll	89	0.31	95	0.21
X_1X_2Y_trained_syll	63	0.48	79	0.41
XXY_untrained_syll	56	0.50	63	0.48
X_1X_2Y_untrained_syll	86	0.34	91	0.29

Table 1. Results per type.

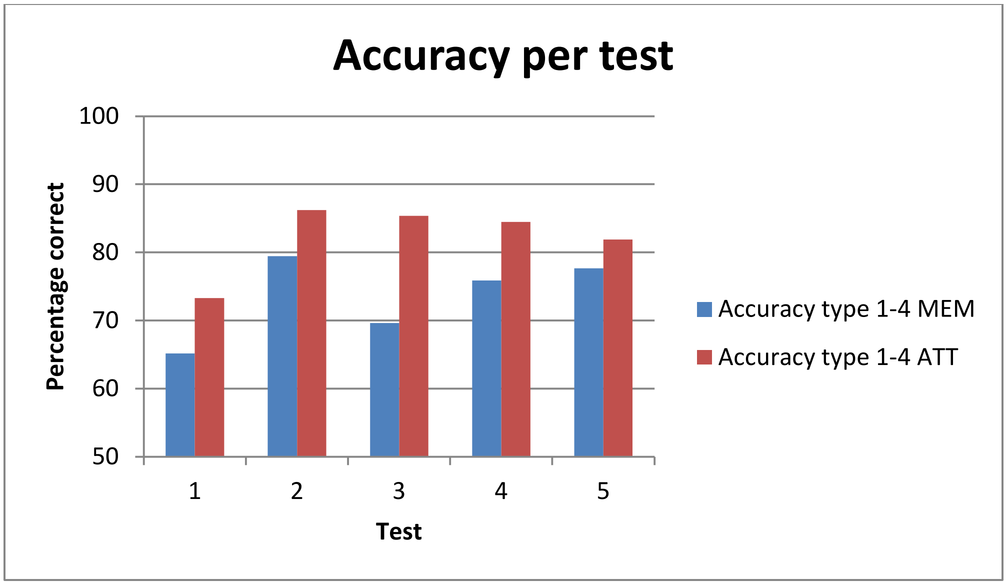

Figure 4. Performance per test (for all types) for the experimental condition (MEM) in blue and the control condition (ATT) in red.

	Memory (MEM)		Attention (ATT)	
	Mean(%)	**SD**	**Mean(%)**	**SD**
Test 1	65	0.48	73	0.44
Test 2	79	0.40	86	0.34
Test 3	70	0.46	85	0.35
Test 4	76	0.43	84	0.36
Test 5	78	0.42	82	0.39

Table 2. Results per test. Average percentage of correct responses for each of the three intermediate test phases and the final test (consisting of test 4 and 5).

To test the effect of group on performance, the experimental group (MEM) and the control group (ATT) were compared in a Generalised Linear Mixed Model in SPSS with accuracy scores of the rule induction task (ACCURACY) as dependent variable. The group the participants were in (GROUP), the test the answers were given in (TEST), the type of the test items (TYPE), the digit span score (DIGITSP), the d' score of the incidental memorisation task (IMT), the Raven's percentiles (RAVENP) and the interactions GROUP*TYPE, GROUP*TEST were included as fixed factors. Participant ID (PPID) and trial ID (TRIALID) were included as random factors. A significance level of 0.05 was used for all statistical tests. There was a statistically significant effect of GROUP ($F(1,1121) = 4.625$, $p = 0.032$), TYPE ($F(3, 1121) = 11.881$, $p = 0.000$) and TEST ($F(4, 1121) = 4.652$, $p = 0.001$). There was no statistically significant effect for GROUP*TYPE, GROUP*TEST, DIGITSP, IMT or RAVENP. See Appendix F for the Fixed Effects table. The analysis for type showed a significant difference for type 2 and 3. The analysis for test showed a significant difference for test 1 from the other tests.

Post-hoc analyses showed that Type 3 accuracy for the memory group was not significantly above chance ($P (t \geq 1.36) = 0.093$) and type 3 accuracy for the attention was significantly above chance ($P (t \geq 3.36) = 0.001$)[11].

Cohen's d analyses revealed that the effect size between the two groups for type 2 was small to medium (d = 0.37), and the effect size between the two groups for type 3 was small (d = 0.16). The effect sizes between type 2 and 4 for the memory and attention group were calculated. The effect size between type 2 and 4 for the memory group was medium to large (d = 0.56) and the effect size between type 2 and 4 for the attention group was small to medium (d = 0.33).

Independent samples t-test was conducted in SPSS comparing the accuracy scores for all types from 20 participants with the highest digit span scores to the 20 participants with the lowest digit span scores. There was a significant difference between the high scoring group (M = 0.82, SD = 0.39) and the low scoring group (M = 0.76, SD = 0.43); $t(790) = -2.00$, $p = 0.045$. See Appendix G for the group statistics and independent samples test table. No significant results were found for the incidental memorisation task or the Raven's progressive matrices, or for each individual type.

Discussion

This study examined the effect of adding a memory overload to rule induction. This was studied using an artificial grammar task based on the task of Radulescu et al. A secondary task was added to overload the memory capacity. The results will be discussed per significant effect first, starting with the effect of type (as that was a replication of the Radulescu et al. study), continuing with condition (as that was the main question of this study) and ending with test order. After the significant main effects, the non-significant independent memory and pattern recognition tasks will be discussed.

Effect of type

The results also showed a main effect of test item type: types 2 and 3 were significantly different from type 1 and 4 for both groups. This can clearly be seen in the graph in Fig. 5. The performance of the attention group on type 3 items (XXY_untrained_syllables) was significantly

[11] Using the formula $\frac{\bar{x}-\mu_0}{S/\sqrt{n}}$ in which $\bar{x}$ is the observed sample mean, μ_0 the H_0 value (in this case 0.5 as the mean is compared to chance), S the standard error and n the sample size.

different from chance, while the memory group's performance was not significantly different from chance. Even though the difference in effect size was only small, this does show that the attention group performed slightly better on type 3 items, and thus seemed to have more knowledge about the XXY pattern of the language. This small effect could possibly be explained by the less explicit knowledge of the language by the memory group. The effect size between the attention and memory group for type 2 items (X_1X_2Y_trained_syllables) was small to medium, indicating that the attention group performed slightly better than the memory group in rejecting type 2 items. This result could also be explained by the fact that the attention group has more knowledge about the language's XXY pattern; they were better in rejecting items that did not follow this pattern. The incorrect acceptance of type 2 items by the memory group can be explained by the existence of a memory trace for the syllables present in both groups; both groups performed worse in rejecting type 2 items than type 4 items, as shown by the medium effect sizes between the type 2 and type 4 accuracy scores for both groups. This shows that both groups have, to some extent, memory traces of the syllables (despite the attempted memory overload in the memory group), as they perform worse in rejecting words with an incorrect pattern but known syllables (type 2, X_1X_2Y_trained_syllables) than in rejecting words with both an incorrect pattern and unknown syllables (type 4, X_1X_2Y_untrained_syllables). Thus, even though the memory secondary task was supposed to overload the memory of the participants, a memory trace still remained. This memory trace could have been learned implicitly. Perhaps the memory secondary task did not fully overload the memory capacity as was intended. The smaller effect size between the two groups' performance for type 3 items (XXY_untrained_syllables) may be caused by the fact that type 3 items only deal with the known pattern of the language, not the known syllables. The existence of the memory traces for the syllables may have prevented the attention group from accepting more type 3 items. This seems to indicate that both memory traces and rule induction were involved in the attention group at the same time.

To summarise, the attention group was better at rejecting type 2 words than the memory group, as they possibly had more knowledge of the language due to their explicit learning of the language, leading to more awareness of the XXY pattern. The performance of the groups for type 3 was much worse and hardly differed between the two groups. This difference in performance between type 2 and 3 words can be explained by a memory trace of the syllables, as can be seen in the difference in performance for type 2 and type 4 items, and more explicit knowledge about the language and its pattern for the attention group.

Effect of group

The results furthermore showed that there was a main effect of condition on accuracy; the attention group performed better than the memory group on all the types of test items. This seems to confirm only hypothesis 1.4, which stated that a secondary task will cause the participant to pay less attention to the language, thus causing her/him to learn the language only implicitly (instead of implicitly and explicitly) and thus leading to less category-based generalisations. The attention group served as a control to the memory group, to make sure that the effect found would be caused by memory and not hearing two sound strings (the language and the secondary task) at the same time while writing numbers down. This could indicate that the memory group had a more demanding, and thus more distracting, secondary task. From the results of this study it would seem that the less explicit attention is paid to the language, the worse the participants perform on all the types, not just type 3 (testing category-based generalisations). To confirm this, however, the results from this study need to be compared to those from the 2.8 bits entropy group in Radulescu et al. This group was exposed to the same language, with the same input entropy of 2.8 bits, but it did not do a secondary task. If the lack of explicit learning modulates the correct acceptance or rejection of test items, this group should perform better than both the attention and memory group. The combined results from this study and the 2.8 entropy group from Radulescu et al. can be seen in Fig. 5 and Table 3.

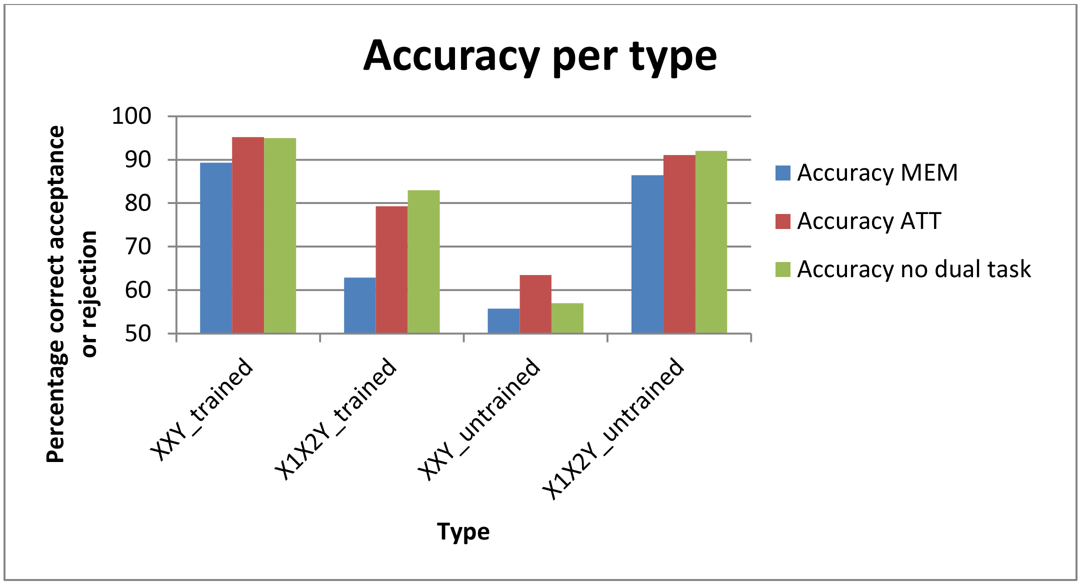

Figure 5. Performance per type for this study and the Radulescu et al. 2.8 entropy group.

	Memory (MEM)		Attention (ATT)		No secondary task (Radulescu et al.)	
	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD
XXY_trained_syll	89	0.31	95	0.21	95	0.22
X_1X_2Y_trained_syll	63	0.48	79	0.41	83	0.50
XXY_untrained_syll	56	0.50	63	0.48	57	0.28
X_1X_2Y_untrained_syll	86	0.34	91	0.29	92	0.37

Table 3. Results per type for this study and the Radulescu et al. 2.8 entropy group.

The overall results from Radulescu et al. show the same pattern as the results of this study: type 1 correct acceptance and type 4 correct rejections are very high, type 2 correct rejections were somewhat lower and type 3 correct acceptance is the lowest. At first glance, the results from Radulescu et al. are for type 1, 2 and 4 nearly the same as for the attention group; only type 3 is similar to the memory group. This would indicate that adding a secondary task can lead to more category-based generalisations. However, this is only applicable for the attention group. This could indicate that as long as the learning mechanism stays the same (being both explicit and implicit), adding a secondary task would indeed lead to an increase of category-based generalisations. This partially confirms hypothesis 1.1, stating that adding a memory secondary task will overload the channel capacity leading to an increase in category-based generalisations. If the secondary task becomes too demanding (as in the memory group), the knowledge about the language is learned more via implicit learning and less via explicit learning. This leads to an overall decrease in performance. Nothing specific can be said about category-based generalisations (acceptance of XXY_untrained_syll words) in the memory group, as this result did not significantly differ from chance.

To summarise, the data suggests that the addition of a secondary task can overload the channel capacity, leading to an increase in category-based generalisations, only if the secondary task is not too distracting and thereby changing the amount of attention paid to the language. Should the secondary task take away too much attention from the language, the language seems to be learned more implicitly and less explicitly, thus leading to less knowledge about the language and an overall worse performance.

Effect of test order

The results showed that there was a main effect of test, showing that test 1 was consistently lower than the other four tests. This showed that the participants learned more about the language over time. However, this result was only significant for test 1 and thus after training phase 2, the learning did not increase significantly.

The brain's channel

From the results from this study it is unclear how many channels there are in the brain. The results for type 3 for the memory group are similar to those of the 2.8 entropy group of Radulescu et al. However, as the memory overload probably led to less awareness and explicit learning, hypotheses 1.2 and 1.3 (stating that a memory overload will not lead to an increase in category-based generalisations) cannot yet be rejected. However, as hypothesis 1.1 (stating that a memory overloading task would lead to an increase in category-based generalisations) was only confirmed for the addition of a secondary task that was not demanding (the attention secondary task) and not the more demanding secondary task (the memory secondary task), the only claim that can be made from these results is that attention seems to underlie the channel capacity. This is not unexpected as it was present in the unified memory model on p. 15 as a factor that directs the working memory. No definitive claim can be made about memory, as the memory secondary task not only affected the channel capacity in terms of memory, but also in terms of amount of attention, leading to a decrease in explicit learning and knowledge of the language. A possible explanation can be that memory does not underlie the brain's channel capacity (as overloading it did not lead to more category-based generalisations), but the overall worse accuracy scores for the memory group for all types seem to indicate that the additional attention load interfered with the memory load. Thus, hypotheses 1.2 and 1.3 are neither rejected nor confirmed. Further research is needed to examine the effect of memory separate from attention, needed to confirm or reject these hypotheses and make more claims about the relation between memory and the channel capacity.

Independent memory and pattern recognition capacities

There was no significant main effect for the independent memory and pattern recognition measures. However, when looking at the difference in accuracy scores between the high and low scoring participants on the three tests, there was a significant difference for the digit span test. A high digit span indicated a higher accuracy score overall. This did not fully reject hypothesis 3, stating a high memory capacity led to less category-based generalisations, as this result was only found for the overall score and not for type 3 (XXY_untrained_syllables) specifically. This seems to indicate that the variation in the participants was too small. If maturation and decline of the brain's channel capacity occur, this small variation could be caused by the small variation in age or the small sample size. Further research testing more participants with more variation in age may lead to significant results in the other two tests as well. It could be that a participant's incidental memory capacity memory and pattern recognition capacity are unrelated to rule induction. This would reject hypothesis 4, stating that a higher pattern recognition capacity leads to more category-based generalisations. However, as the difference between explicit and implicit learning do seem to underlie the difference between the groups, this does not seem to be the case for incidental memorisation. A second reason why there was no main effect for either the digit span task and the incidental memorisation task could be that they test the short term memory only. Rule learning is also concerned with the order of the syllables and this is seen by some as a form of manipulation. Perhaps the OPSAN test used by De Jong (2010) may be better to test the effect of working memory on rule induction. As languages are expected to have patterns in the form of rules, the pattern recognition capacity is still expected to underlie rule induction as previously found by Little et al. (2012). However, the fact that the task was visual could have affected the results. Even though Altmann et al. (1995) found that the ability to generalise rules could transfer to other domains, perhaps this is only true for learning the same rule in vision as sound and not a general capacity. Throughout the entire memory model in Fig. 2, vision and sound are separate: from the visual and phonological input, to the iconic and echoic memory in the sensory memory, to the phonological loop and the visuo-spatial sketchpad to the linguistic memory and visual knowledge in the long term memory. Perhaps an auditory pattern recognition task may better represent the auditory pattern recognition capacity.

To summarise, the lack of main effect for the independent memory and pattern recognition capacities may be caused by the lack of variation in participants' age or the small sample size. As implicit, or incidental, memorisation and pattern recognition capacities are still expected to influence rule induction, perhaps other tests may be more suitable in the future in a small group with less variation in age as used in this study.

Conclusion

The aim of this paper was to find out if overloading the brain's channel capacity by means of a memory secondary task led to an increase in category-based generalisations. The results of this study indicate that the memory secondary task used was too distracting. It diverted attention away from the language and this caused the artificial language to be learned more implicitly than explicitly. This lead to a decrease in overall performance. The secondary task in which only attention was overloaded, using a less demanding secondary task, did seem to lead to an increase in category-based generalisations compared to the original group without any secondary task from Radulescu et al. This shows that attention underlies the brain's channel capacity. Due to the difference in learning mechanisms (implicit versus explicit learning) between the two groups, nothing certain can be said about the role of memory in the brain's capacity. Future research using different secondary tasks (distracting the participant equally) may shed more light on this. A second line of research was whether individual memory and pattern recognition capacities influence a participant's ability to make category-based generalisations. However, the results showed no significant main effects and this may have been caused by a lack of variation in participants' ages or the small sample size.

Another obvious line of future research is to study the effect of a pattern recognition secondary task on rule induction hypothesised about in this paper. When designing such an experiment, the amount of attention must be taken into consideration. The results from this study concerning explicit and implicit learning must be taken into consideration for educational purposes, as they show that the additive effect of explicit learning lead to a significantly better performance overall. Lastly, to further understand the role of a secondary task on rule induction, neuro-imaging techniques can be used to track online activation in the brain (using for example the temporally good EEG and MEG techniques) to study how the brain is involved in rule-induction and memorisation and perhaps find physical evidence for the brain's channel.

References

Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 899.

Anderson, J. R. (1990). *Cognitive psychology and its implications* (8th ed.). New York, NY: Worth Publishers.

Baars, B. J., & Gage, N. M. (2007). *Cognition, brain, and consciousness: Introduction to cognitive neuroscience* (2nd ed.). Amsterdam: Academic Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, 4(11), 417-423.

Baddeley, A., Eysenck, M. W, & Anderson, M. C. (2015). *Memory* (2nd ed.). London: Psychology Press.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in cognitive sciences*, *2*(10), 406-416.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science*, 38(2), 105-134.

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *Academy of management review*, *10*(4), 803-813.

Goldsmith, J. (2000). On information theory, entropy, and phonology in the 20th century. *Folıa lınguıstıca*, *34*(1-2), 85-100.

Goldsmith, J. (2007). Probability for linguists. *Mathématıques et scıences humaınes. Mathematıcs and socıal scıences*, (180), 73-98.

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping timesharing, II: Auditory secondary task. *Acta Psychologica*, *40*(5), 343-362.

Knowlton, B. J. & Squire, L. R. (1994). The information acquired during artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn. 20*, 79–91.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, *14*(4), 389-433.

Lachman, S. J. (1997). Learning is a process: Toward an improved definition of learning. *The Journal of psychology*, *131*(5), 477-480.

Lien, M. C., Ruthruff, E., & Johnston, J. C. (2006). Attentional limitations in doing two tasks at once: The search for exceptions. *Current Directions in Psychological Science*, *15*(2), 89-93.

Little, D. R., Lewandowsky, S., & Griffiths, T. (2012). A Bayesian model of rule induction in Raven's progressive matrices. In *Proceedings of the 34th annual conference of the cognitive science society* (pp. 1918-1923).

López-Barroso, D., Cucurell, D., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2016). Attentional effects on rule extraction and consolidation from speech. *Cognition*, *152*, 61-69.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

McAllister, T. W., Sparling, M. B., Flashman, L. A., Guerin, S. J., Mamourian, A. C., & Saykin, A. J. (2001). Differential working memory load effects after mild traumatic brain injury. *Neuroimage*, *14*(5), 1004-1012.

McDowell, S., Whyte, J., & D'Esposito, M. (1997). Working memory impairments in traumatic brain injury: evidence from a dual-task paradigm. *Neuropsychologia*, *35*(10), 1341-1353.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science*, *291*(5501), 114-118.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Pashler, H. (Ed.). (2016). *Attention*. Psychology Press.

Perruchet, P. & Pacteau, C. (1990). Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge? *J. Exp. Psychol. General*, *119*, 264–275.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological bulletin*, *133*(2), 227.

Pothos, E. M. (2010). An entropy model for Artificial Grammar Learning. *Frontiers in Cognitive Science, 1*, 1-13.

Radulescu, S., Wijnen, F. N. K., & Avrutin, S. (in prep). Patterns bit by bit. An Entropy Model for Linguistic Generalizations.

Raven Standard Progressive Matrices (SPM). (2006). (Elycio (v/h Elesevier) Vertaalbureau, Trans). Bloomington, MN: NCS Pearson, Inc. (Original work published 2004).

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of experimental psychology:*

General, *118*(3), 219.

Reber, A. S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition*, *5*(4), 333-361.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of experimental psychology: general*, *125*(2), 123.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technıcal Journal* 27, 379-423.

Stark, C. E., & Okado, Y. (2003). Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. *Journal of Neuroscience, 23*(17), 6748-6753.

Van Ewijk, L., & Avrutin, S. (2016). Lexical access in non-fluent aphasia: a bit more on reduced processing. *Aphasiology*, *30*(11), 1264-1282.

Wills, A., Barrasin, T., & McLaren, I. (2011, January). Working memory capacity and generalization in predictive learning. In *Proceedings of the Cognitive Science Society* (Vol. 33, No. 33).

Yang, H., Yang, S., & Park, G. (2013). Her voice lingers on and her memory is strategic: Effects of gender on directed forgetting. *PloS one*, *8*(5), e64030.

Appendices

Appendix A

Illustration of entropy

To explain the entropy formula, let us first look at a situation in which the probability of all the variables is (theoretically) equal. This would be the case with a coin flip; there are two sides (or values), with an equal probability of both sides being thrown. The probability of one of the sides would be ½ or 0.50, as there is 50% chance of the coin flip going either way. Filling out the entropy formula would give:

$$-(½*log_2(½)+½*log_2(½))$$
$$= -log_2(½)$$
$$= 1$$

A second example of a situation in which the probability of all the variables is (theoretically) equal, is the case of a dice throw. There are six sides (or values) to a die, with an equal probability of one of the six sides being thrown. This probability would be 1/6. Filling out the entropy formula with these numbers would give:

$$-(1/6*log_2(1/6)+1/6*log_2(1/6)+1/6*log_2(1/6)+1/6*log_2(1/6)+1/6*log_2(1/6)+1/6*log_2(1/6))$$
$$= -log_2(1/6)$$
$$\approx 2.58$$

These two examples show that increasing the number of variables or options, even when the probabilities of all the variables are the same, increases the entropy. This makes sense if Goldsmith's description of entropy is kept in mind: “entropy is a measure of the unpredictable character of a set of objects” (2000, p. 86). If the number of options increases, it becomes less predictable what the outcome will be and thus, the entropy increases as the uncertainty about the outcome increases. Keeping this definition in mind, it can be predicted that if the probability of one of the variables increases, there will be less uncertainty to the outcome (as there is a bigger chance that one of the variables will be the outcome). Thus, the entropy will decrease. To illustrate this, let's imagine a die in which the 6-side is weighted (thus increasing the chance that the 1-side will come up). This biased die has a 50% (or 0.50) chance of one being the outcome.

Let's further imagine that all the other numbers have an equal probability of being thrown (thus 10% or 0.10 per side). Filling out the entropy formula with this information would give:

$$-(0.5*log_2(0.5)+0.1*log_2(0.1)+0.1*log_2(0.1)+0.1*log_2(0.1)+0.1*log_2(0.1)+0.1*log_2(0.1))$$
$$= -(0.5*log_2(0.5)+0.5*log_2(0.1)$$
$$\approx 2.16$$

As can be seen, if the probability of all the variables is not even, there is a higher chance of the outcome being one of the items in the set, leading to less uncertainty and thus less entropy. If there are several items in a set all with equal probability, then there is more uncertainty about the outcome and thus the entropy is highest.

Appendix B

Syllables and calculations used in Radulescu et al. (in prep) experiment 1

Low entropy condition

X-syllables	Y-syllables
4*goo	4*sjie
4*puu	4*duu
4*teu	4*saa
4*vee	4*feu
4*woo	4*moo
4*loo	4*kee

H[bX] = H[6] = $-log_2(0.167)$ = 2.58
H[XX] = H[6] = $-log_2(0.167)$ = 2.58
H[XY] = H[24] = $-log_2(0.042)$ = 4.58
H[Ye] = H[6] = $-log_2(0.167)$ = 2.58
H[bXX] = H[6] = $-log_2(0.167)$ = 2.58
H[XXY] = H[24] = $-log_2(0.042)$ = 4.58
H[XYe] = H[24] = $-log_2(0.042)$ = 4.58
H[bigram] = (2.58+2.58+4.58+2.58)/4 = 3.08
H[trigram] = (2.58+4.58+4.58)/3 = 3.91
H[total] = (3.08+3.91)/2 = 3.5

Medium entropy condition

X-syllables		Y-syllables	
2*goo	2*goe	2*sjie	2*sjoe
2*puu	2*heu	2*duu	2*weu
2*teu	2*juu	2*saa	2*fie
2*vee	2*nie	2*feu	2*kaa
2*woo	2*roo	2*moo	2*muu
2*loo	2*vuu	2*kee	2*choo

H[bX] = H[12] = $-log_2(0.083)$ = 3.58
H[XX] = H[12] = $-log_2(0.083)$ = 3.58
H[XY] = H[24] = $-log_2(0.042)$ = 4.58
H[Ye] = H[12] = $-log_2(0.083)$ = 3.58
H[bXX] = H[12] = $-log_2(0.083)$ = 3.58
H[XXY] = H[24] = $-log_2(0.042)$ = 4.58
H[XYe] = H[24] = $-log_2(0.042)$ = 4.58
H[bigram] = (3.58+3.58+4.58+3.58)/4 = 3.83
H[trigram] = (3.58+4.58+4.58)/3 = 4.25
H[total] = (3.83+4.25)/2 = 4

High entropy condition

X-syllables				Y-syllables			
goo	goe	haa	luu	sjie	sjoe	woe	meu

puu	heu	hie	noo	duu	weu	sjeu	chie
teu	juu	jie	noe	saa	fie	buu	Zuu
vee	nie	joe	ruu	feu	kaa	daa	fuu
woo	roo	jeu	vie	moo	muu	faa	boo
loo	vuu	lie	veu	kee	choo	koo	huu

H[bX]	= H[12]	$= -log_2(0.042)$	= 4.58
H[XX]	= H[12]	$= -log_2(0.042)$	= 4.58
H[XY]	= H[24]	$= -log_2(0.042)$	= 4.58
H[Ye]	= H[12]	$= -log_2(0.042)$	= 4.58
H[bXX]	= H[12]	$= -log_2(0.083)$	= 3.58
H[XXY]	= H[24]	$= -log_2(0.042)$	= 4.58
H[XYe]	= H[24]	$= -log_2(0.042)$	= 4.58
H[bigram]	= (4.58+4.58+4.58+4.58)/4		= 4.58
H[trigram]	= (4.58+4.58+4.58)/3		= 4.58
H[total]	= (4.58+4.58)/2		= 4.58

Syllables and calculations used in Radulescu et al. (in prep) experiment 2

Low entropy condition

X-syllables	Y-syllables
4*kee	4*muu
4*joe	4*goo
4*daa	4*lie
4*puu	4*vee
4*teu	4*reu
4*hie	4*saa
4*foo	4*sjoe

H[bX]	= H[7]	$= -log_2(0.143)$	= 2.8
H[XX]	= H[7]	$= -log_2(0.143)$	= 2.8
H[XY]	= H[7]	$= -log_2(0.143)$	= 2.8
H[Ye]	= H[7]	$= -log_2(0.143)$	= 2.8
H[bXX]	= H[7]	$= -log_2(0.143)$	= 2.8
H[XXY]	= H[7]	$= -log_2(0.143)$	= 2.8
H[XYe]	= H[7]	$= -log_2(0.143)$	= 2.8
H[bigram]	= (2.8+2.8+2.8+2.8)/4		= 2.8
H[trigram]	= (2.8+2.8+2.8)/3		= 2.8
H[total]	= (2.8+2.8)/2		= 2.8

Medium entropy condition

X-syllables		Y-syllables	
2*kee	2*noo	2*muu	2*sjeu
2*joe	2*noe	2*goo	2*veu
2*daa	2*kuu	2*lie	2*waa
2*puu	2*jeu	2*vee	2*vie
2*teu	2*too	2*reu	2*meu

2*hie	2*haa	2*saa	2*vuu
2*foo	2*fuu	2*sjoe	2*sjie

H[bX]	= H[14]	$= -log_2(0.071)$	= 3.8
H[XX]	= H[14]	$= -log_2(0.071)$	= 3.8
H[XY]	= H[28]	$= -log_2(0.036)$	= 4.8
H[Ye]	= H[14]	$= -log_2(0.071)$	= 3.8
H[bXX]	= H[14]	$= -log_2(0.071)$	= 3.8
H[XXY]	= H[28]	$= -log_2(0.036)$	= 4.8
H[XYe]	= H[28]	$= -log_2(0.036)$	= 4.8
H[bigram]	= (3.8+3.8+4.8+3.8)/4		= 4.05
H[trigram]	= (3.8+4.8+4.8)/3		= 4.47
H[total]	= (4.05+4.47)/2		= 4.26

High entropy condition

X-syllables				Y-syllables			
kee	noo	doo	kaa	muu	sjeu	geu	zeu
joe	noe	buu	keu	goo	veu	roo	loe
daa	kuu	bie	boo	lie	waa	moo	gee
puu	jeu	kie	dee	vee	vie	goe	vaa
teu	too	fie	huu	reu	meu	zuu	seu
hie	haa	foe	faa	saa	vuu	weu	luu
foo	fuu	heu	juu	sjoe	sjie	woo	chie

H[bX]	= H[28]	$= -log_2(0.036)$	= 4.8
H[XX]	= H[28]	$= -log_2(0.036)$	= 4.8
H[XY]	= H[28]	$= -log_2(0.036)$	= 4.8
H[Ye]	= H[28]	$= -log_2(0.036)$	= 4.8
H[bXX]	= H[28]	$= -log_2(0.036)$	= 4.8
H[XXY]	= H[28]	$= -log_2(0.036)$	= 4.8
H[XYe]	= H[28]	$= -log_2(0.036)$	= 4.8
H[bigram]	= (4.8+4.8+4.8+4.8)/4		= 4.8
H[trigram]	= (4.8+4.8+4.8)/3		= 4.8
H[total]	= (4.8+4.8)/2		= 4.8

Appendix C

Stimuli used in the rule induction task and the secondary tasks

Stimuli used in the rule induction task

Training items

1 Kee-kee-muu	8 Kee-kee-muu	15 Kee-kee-muu	22 Kee-kee-muu
2 Joe-joe-goo	9 Joe-joe-goo	16 Joe-joe-goo	23 Joe-joe-goo
3 Daa-daa-lie	10 Daa-daa-lie	17 Daa-daa-lie	24 Daa-daa-lie
4 Puu-puu-vee	11 Puu-puu-vee	18 Puu-puu-vee	25 Puu-puu-vee
5 Teu-teu-reu	12 Teu-teu-reu	19 Teu-teu-reu	26 Teu-teu-reu
6 Hie-hie-saa	13 Hie-hie-saa	20 Hie-hie-saa	27 Hie-hie-saa
7 Foo-foo-sjoe	14 Foo-foo-sjoe	21 Foo-foo-sjoe	28 Foo-foo-sjoe

Test items

Type 1	**Type 2**	**Type 3**	**Type 4**
Daa-daa-lie	Joe-daa-saa	Duu-duu-taa	Poo-gaa-roe
Hie-hie-saa	Peu-teu-muu	Zoe-zoe-voo	Roe-nuu-nie
Kee-kee-muu	Kee-foo-vee	Soo-soo-ruu	Gaa-mie-suu
Teu-teu-reu	Hie-daa-reu	Jie-jie-feu	Suu-nie-nuu
Joe-joe-goo	Teu-puu-goo	Woe-woe-see	Mie-poo-gaa

Stimuli used in the secondary tasks

Memory condition							
Warm up		Training phase 1		Training phase 2		Training phase 3	
Input	Correct answer	Input	Correct answer	Input	Correct answer	Input	Correct answer
4		9		3		7	
3		6		5		2	
7		1		1		4	
*[12]	3	*	6	2		*	2
5		2		*	1	5	
8		7		7		3	
*	5	*	2	0		8	
2		4		8		9	
0		3		*	0	*	8
9		8		6		1	
6		5		1		6	
*	9	*	8	*	6	*	1
1		7		4		0	
7		0		9		4	
4		9		*	4	7	
		4		5		9	
		*	9	7		*	7

[12] The asterisk (*) is used to indicate the "bleep" sound.

		2		3		2	
		3		0		5	
		8		*	3	3	
		*	3	2		*	5
		2		5		8	
		5		*	2	0	
		0		1		*	8
		*	5	8		4	
		7		4		8	
		3		*	8	2	
		*	7	6		3	
		8		3		*	2
		6		9		6	
		9		4		1	
		3		*	9	5	
		*	9	1		*	1
		4		7		7	
		2		*	1	5	
		1		5			

Attention condition							
Warm up		Training phase 1		Training phase 2		Training phase 3	
Input	Correct answer	Input	Correct answer	Input	Correct answer	Input	Correct answer
*		*		*		*	
*		*		*		*	
*		4	4	*		4	4
4	4	*		8	8	*	
*		*		*		*	
*		1	1	*		*	
9	9	*		*		*	
*		*		5	5	7	7
*		*		*		*	
*		0	0	*		*	
*		*		*		*	
3	3	*		*		3	3
*		*		4	4	*	
*		*		*		*	
1	1	5	5	*		1	1
*		*		*		*	
*		*		2	2	*	
*		6	6	*		9	9
		*		*		*	
		*		*		*	
		*		3	3	*	
		2	2	*		5	5
		*		*		*	

		*		7	7	*	
		9	9	*		*	
		*		*		*	
		*		9	9	8	8
		*		*		*	
		7	7	*		*	
		*		*		*	
		*		*		*	
		*		6	6	2	2
		*		*		*	
		3	3	*		*	
		*		1	1	*	
		*		*		6	6
		*		*		*	
		5	5	*		*	
		*		8	8	7	7
		*		*		*	

Appendix D

Stimuli used in the incidental memorisation task

Training words from the categorisation phase	Test items in the incidental memorisation task	Correct response for the incidental memorisation task
Baa-duk	Blie-ker	Yes
Blie-ker	Die-taa	Yes
Die-taa	Faa-poeg	Yes
Doo-moo	Grie-fup	Yes
Faa-poeg	Kaa-sie	Yes
Floe-nie	Kle-pin	Yes
Grie-fup	Loo-ga	Yes
Kaa-sie	Maa-lon	Yes
Kier-tan	Nil-boo	Yes
Kle-pin	Per-gon	Yes
Loo-ga	Rie-wan	Yes
Maa-lon	Stee-poer	Yes
Na-spu	Vaa-mie	Yes
Nil-boo	Ro-gges	Yes
Per-gon	Waa-bo	Yes
Pli-zet	Ti-pla	No
Ra-jee	Baa-nip	No
Rie-wan	Duu-foo	No
See-ta	Fie-dang	No
Sie-long	Goo-pem	No
Stee-poer	Ken-gel	No
Tar-sin	Loo-tup	No
Vaa-mie	Mui-bloo	No
Wi-ffel	Nij-foe	No
Waa-bo	Noe-baa	No
Den-sim	Ra-stin	No
Hi-ftam	Ree-ming	No
Kij-bog	Sei-bor	No
Moo-vig	Vei-rig	No
Ro-gges	Soe-lep	No

Appendix E

Example of a Raven's Standard Progressive Matrix

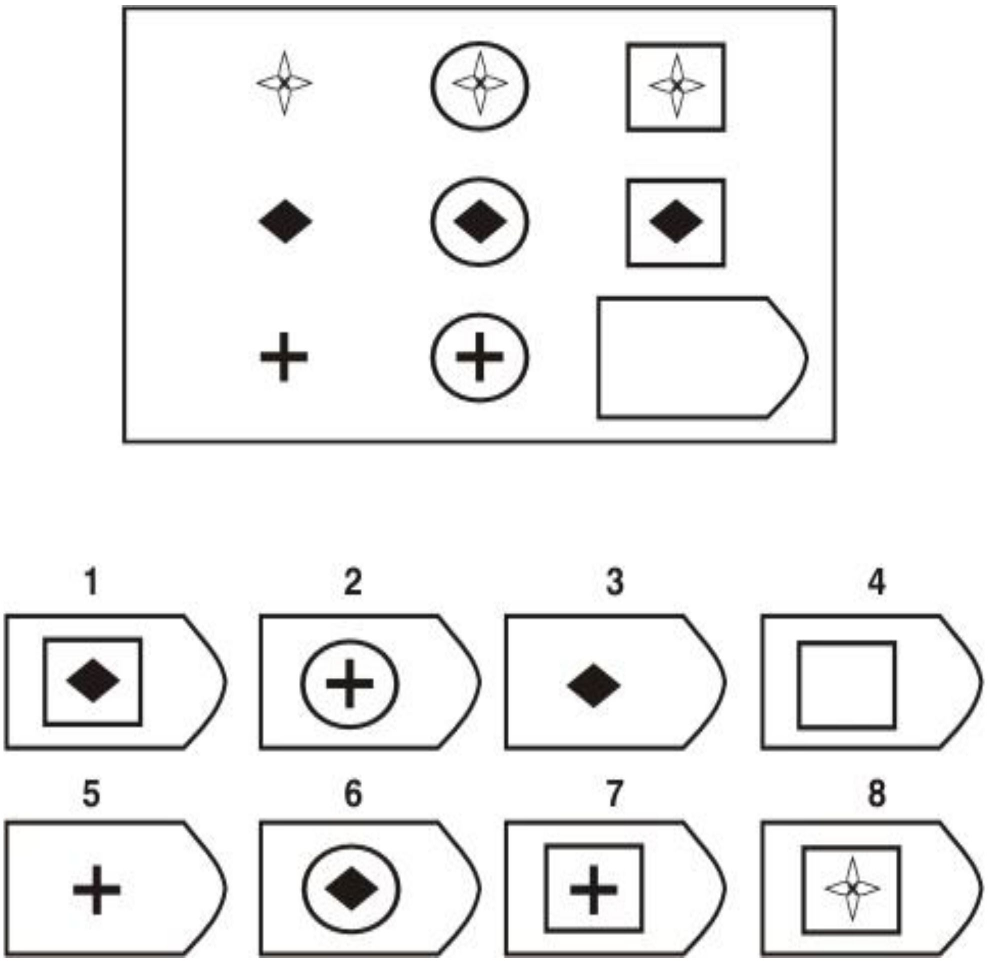

Figure 6. Example of a Raven's Standard Progressive Matrix (D4).

Appendix F

Fixed effects table of results and descriptive statistics of random effects

Fixed Effects

Target:ACCURACY
Reference Category:1

Source	F	df1	df2	Sig.
Corrected Model ▼	3,476	18	1.121	,000
GROUP	4,625	1	1.121	,032
TYPE	11,881	3	1.121	,000
TEST	4,652	4	1.121	,001
GROUP*TEST	0,842	4	1.121	,498
GROUP*TYPE	1,082	3	1.121	,356
DIGITSP	2,042	1	1.121	,153
IMT	0,179	1	1.121	,672
RAVENP	0,497	1	1.121	,481

Probability distribution:Binomial
Link function:Logit

Figure 7. Fixed effects table of results.

Appendix G

Results digit span scores on accuracy

	DigitGroup	N	Mean	Std. Deviation	Std. Error Mean
ACCURACY	1	400	,76	,426	,021
	2	400	,82	,385	,019

Table 4. Group Statistics. Group 1 consisted of the 20 participants with the lowest digit spans, group 2 consisted of the 20 participants with the highest digit spans.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ACCURACY	Equal variances assumed	16,222	,000	-2,003	798	,045	-,057	,029	-,114	-,001
	Equal variances not assumed			-2,003	789,799	,045	-,057	,029	-,114	-,001

Table 5. Independent Samples Test.