Author: Geambasu, A.
Title: Simple rule learning is not simple : studies on infant and adult pattern perception and production
Issue Date: 2018-12-11

CHAPTER 1

Introduction

Human language is an acquired symbolic system of communication and thought. When acquiring language, (hearing) infants must learn a great deal, including the sound inventory of their language, how those sound categories may be combined in their native language to make words, and how the words may be combined into well-formed phrases. Along the way, they are also simultaneously learning the meanings and referents of these words and utterances. How infants are able to learn this complex system has been the focus of much research throughout the history of (psycho-)linguistics and developmental psychology going back centuries (W. Levelt, 2013). How *do* infants make the mental leap from their specific input to the ability to understand and produce an infinite number of utterances? What kinds of cognitive mechanisms are involved in learning the structures underlying language and are those mental mechanisms uniquely suited for learning language, or are they shared with other mental capacities? These are questions that will be addressed in this thesis using the tool of artificial grammar learning tasks.

The work presented here focuses on better understanding the biases infants and adults may have when learning rules, in both perception and production, whether these biases are domain-general or specific to language, how they may change in adulthood, and how they may interact with task-specific factors and attention. The experiments were developed as part of a larger project studying similar biases and learning mechanisms in birds (Spierings, 2016) and devising computational models of rule learning (Alhama, 2017).

The experimental work in this project was conducted by means of artificial grammar learning tasks. I will first introduce the motivation behind using artificial grammar learning tasks, particularly with infants. A detailed outline of the thesis will follow, summarizing the biases and specificity of rule learning that will be addressed in each chapter.

1.1 Artificial grammar learning

At the most basic level, the task infants are facing when acquiring their native language involves discovering both discrete units, like phonemes and words, and the rules that regulate the ways in which these units can be combined. Despite what this highly simplified definition of language acquisition may suggest, it is a daunting task. Nevertheless, infants are highly sensitive to linguistic input, and already begin to perceive and process aspects of speech and language in the womb (Moon, Cooper, & Fifer, 1993). Speech perception experiments have revealed that soon after birth, infants already have the prosodic knowledge to discriminate their mother's language from languages with a different prosodic structure (Christophe & Morton, 1998), and can soon use statistical information in language to discover discrete units such as phoneme categories (Maye, Werker, & Gerken, 2002). How humans, particularly young infants, are subsequently able to form larger combinatorial units, such as words and sentences, from these discrete items is one of the main questions of linguistic research.

Over the past fifty years, beginning with work from Reber (1967), artificial language learning (ALL) or artificial grammar learning (AGL), paradigms have been important tools for studying our ability to extract discrete units from the speech stream, and to discover the rules combining them. The grammars used in AGL tasks may vary in their complexity and naturalness, but both the discrete units of which they are composed, and their structure, are typically selected such that potential confounds that can lead to solving the task in ways unrelated to the research question – which may appear in natural languages – are minimized. As such, AGL paradigms have been crucial in discovering learning abilities and limitations on learning across age groups, across species, and across domains (e.g., Knowlton & Squire, 1996; Perruchet & Pacteau, 1990; Gómez, 1997).

Following a seminal study from Saffran, Aslin, and Newport (1996) showing that eight-month-old infants are able to calculate statistical regularities in a continuous speech stream to extract words, the first AGL experiment studying infants' ability to learn and generalize rules to novel instances soon followed (Marcus, Vijayan, Bandi Rao, & Vishton, 1999). Using the Headturn Preference Procedure (Kemler Nelson et al., 1995), they exposed seven-month-old infants to a two-minute stream of synthetic speech composed of syllables arranged in an XYX, XYY, or XXY exposure grammar. During a subsequent testing phase in which novel syllables either followed the exposure grammar or a grammar inconsistent with the exposure, the infants showed sensitivity to the difference between the two grammars by longer looking times for the inconsistent test stimuli. The authors thus concluded that in addition to a statistical learning mechanism for learning fixed units, infants have an algebraic rule learning mechanism, where "algebraic" refers to the ability to extrapolate X and Y variables from the familiarization stimuli, allowing for combinations of novel syllables as long as they obey the combinatorial rule as a whole.

Marcus et al.'s (1999) work was seminal in the field of infant AGL, paving the way for novel experiments using variations on this testing paradigm and using different types of stimuli. Subsequent work focused on how these algebraic representations are formed (e.g., Gerken, 2006; Gerken & Bollt, 2008; Gerken, 2010), and on whether this type of learning was domain-specific (as claimed in Marcus, Fernandes, & Johnson, 2007) or domain-general (as claimed in Saffran, Pollak, Seibel, & Shkolnik, 2007). However, nearly two decades after the original publication of Marcus et al.'s initial paper, our understanding of the ability to learn these abstract rules is still lacking, and basic questions about perceptual biases that might influence learning, the domain-specificity of the process, and factors that might facilitate or hinder the learning process are still not answered adequately. Many extensions on the Marcus et al. paradigm have appeared in the literature, offering almost as many different accounts of the type of stimuli that facilitate, or lead to, learning, e.g., speech (Gerken, 2006, 2010; Gerken, Dawson, Chatila, & Tenenbaum, 2015), familiar stimuli (Saffran et al., 2007), stimuli with a communicative function (Ferguson & Lew-Williams, 2016), redundant stimuli (Thiessen, 2012), multimodal stimuli (Frank, Slemmer, Marcus, & Johnson, 2009), surprising stimuli (Gerken et al., 2015). However, despite several notable extensions on the work, the original findings were never replicated exactly. In fact, the literature gives us reason to believe that exact replication of this work is quite difficult. Most concretely, a footnote in Gerken's (2006) paper states that replication with the same age group and same grammars failed. The fact that there are no successful exact replications of the original work motivated the work conducted in the chapter 2 of this thesis. This was initially aimed at an extension of the Marcus study, but when this failed, we attempted to replicate the study as closely as possible. When this replication attempt failed, the focus of the thesis shifted to the study of a variety of factors that may influence (primarily) Marcus-type rule-learning abilities in infants as well as adults. The specific factors and themes that will be examined in this thesis are presented below, and the chapters in which these factors feature will be indicated.

1.2 Domain specificity

One of the goals of this thesis is to understand whether the ability to learn and generalize simple patterns in linguistic structure is part of a domain general learning capacity. Several studies have tested infants with Marcus grammars composed of non-speech stimuli. For instance Marcus et al. (2007) expanded on the 1999 rule learning work in a follow-up study showing that 7.5-month-old infants exposed to either the XYY or XYX pattern in stimuli consisting of tones, timbres, or animal sounds were not able to learn the rule and generalize it to novel input in the same domain. However, if the infants were exposed to sung syllables, they could generalize the rule to novel sung syllables. In addition, the infants could transfer a rule learned on the basis of speech stimuli to test stimuli consisting of non-linguistic tones, timbres, and animal sounds. These findings indicate that rule learning is facilitated by speech, and that rules learned with speech stimuli can be transferred to non-linguistic stimuli. The authors thus argue that there may be a special role for speech in the ability to learn algebraic rules.

Others, however, have shown evidence of rule learning in the visual domain as well, albeit with limitations. Johnson, Fernandes and Frank (2009) tested eight and 11-month-olds on visual stimuli consisting of sequentially-presented triads of shapes. Although some learning was possible in this domain, Johnson et al. found limitations on generalization depending on whether the familiarization grammar contained an immediate repetition at the beginning (XXY), at the end (XYY), or not at all (XYX). The eight-month-olds could not discriminate the grammars when trained on triads with an early repetition (XXY) or a non-adjacent repetition (XYX). This age group only succeeded when they were trained on triads with a late repetition (XYY) and tested with inconsistent XYX stimuli. Infants at the age of 11 months could learn from familiarization stimuli with both early and late repetitions (XXY and XYY), but could still not learn from XYX familiarization stimuli. The early learning bias for late repetition patterns in visual stimuli is also found in Rabagliati, Senghas, Johnson, and Marcus (2012), where the stimuli were videos of sign language signs. The authors found that infants at 7.5 months old were able to generalize XYY but not XXY rules to novel input. Saffran et al. (2007), however, showed that infants could learn from both XYY and XYX visual stimuli if familiarization triads consisted of pictures of dogs or cats.

While rule learning in infants thus seems to be possible in non-linguistic domains, there appear to be limitations that need to be further explored, relating to repetition and the role of stimulus familiarity.

1.3 Repetition

Perception

One of the primary factors that seem to play a role in rule learning is repetition. A perceptual sensitivity to the presence of repetition in stimuli has been found across age groups (e.g., Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Gómez & Gerken, 1999; Endress, Scholl, & Mehler, 2005), species (e.g., Chen, van Rossum, & ten Cate, 2015; van Heijningen, Chen, van Laatum, van der Hulst, & ten Cate, 2013), and domains (e.g., S. P. Johnson et al., 2009). Endress and colleagues (Endress & Bonatti, 2007; Endress, Nespor, & Mehler, 2009; Endress, 2013) have characterized this sensitivity as one of the principal factors aiding in the acquisition of simple statistics and rules during acquisition, which they have labeled Perceptual or Memory Primitives (POMPs). Infants in both speech and visual learning experiments seem more able to learn patterns involving immediate repetition than other patterns. Above, we discussed the visual rule learning work of Johnson et al. where infants could only learn from familiarization stimuli in XYY (8-month-olds) and XXY (11-month-olds) repetition patterns, but not from a non-adjacent repetition pattern XYX. The sensitivity to repetition patterns appears to be present already at birth: in a study using Near-infrared spectroscopy (NIRS), newborns showed more brain oxygenation in response to stimuli organized in the XYY pattern than to stimuli organized in a control pattern XYZ, while they did not show such a differential activation when the pattern XYX was pitted against XYZ (Gervain et al., 2008). It thus seems that, although these Marcus-type grammars appear at first glance equally simple, there is a hierarchy of learnability with respect to these Marcus-type rules, with XYY being the simplest, followed by XXY and XYX.

Learning of rules that do not involve immediate repetition may require extra support. Frank et al. (2009) show that five month old infants can learn both XYX and XYY rules when they are presented with congruent speech and visual information. They could not generalize the rules when they were instantiated only in the speech or in the visual domains. In this case, not only is the visual rule learning aided by speech, but rule learning in speech at this very early age is made possible by the presence of congruent visual information. Although the repetition bias can be overcome given the right input and learning circumstances, it nevertheless remains a strong perceptual bias throughout life. Adults in Marcus grammar learning tasks (Endress, Dehaene-Lambertz, & Mehler, 2007; Chen et al., 2015) and in more complex rule learning paradigms fall back on this POMP to help them generalize to novel input (Endress et al., 2005). The role of repetition is explored in chapters 2 and 3 of this thesis.

Production

Because the evidence for a bias for repetitions in perception is robust, both in the work discussed above and in our own work presented in chapters 2 and 3 of this thesis, we wanted to examine to what extent repetition is also present in production. The role of production abilities in perception preferences, and vice versa, is often neglected, but recent work shows interactions between the two (e.g., DePaolis, Vihman, & Nakai, 2013; Majorano, Vihman, & DePaolis, 2014; ter Haar & Levelt, 2018). To this end, a corpus analysis on babbling data of infants in their first year of life was conducted. Infants begin babbling at around the age of six months. The leading theories of the development of infant babbling state that between this age and the production of their first real words, infants go through a period of development in which they are thought to proceed from mostly repetitive (reduplicated) to more varied (variegated) syllable productions (Stark, 1980; Oller, Yeni-Komshian, Kavanaugh, & Ferguson, 1980; Roug, Landberg, & Lundberg, 1989). Others have identified the entirety of this period as one of reduplication (Koopmans-Van Beinum & van der Stelt, 1986). There seems to be a general consensus with respect to the fact that certain sounds are produced and combined earlier than others by infants, and with respect to a rough time-course of when that happens. However, the data on which these stages have been based have come from a few babies, most of whom were acquiring English, and large-scale corpus analyses to back up these proposed stages are missing (cf. Lipkind et al., 2013).

With our study of the corpus of infant speech available in the CHILDES database, presented in chapter 4, we investigate these developmental stages in longitudinal babbling data from infants from eight different languages. Our aims were both to frame our perceptual findings within the context of what is going on in infants' production during the same period, and to check whether the proposed stages were generalizable to a wider range of languages. Results from the analysis of two-, three-, and four-syllable utterances show asymmetries between patterns including and not including immediate repetitions, with patterns including non-adjacent repetitions (XYX) being very infrequent.

1.4 Familiarity

While it is evident that repetition is one of the key factors governing how and whether infants learn and generalize rules, we also know that infants are able to learn far more complex rules than only those containing immediate repetitions. Other factors should therefore also be at play. Some authors have hypothesized that speech is not special *per se*, but that infants learn rules best on the basis of what is familiar to them (Saffran et al., 2007), and that they invoke rules to make stimuli less "surprising" (Gerken, 2010; Gerken et al., 2015). But the term "familiarity" is not well-defined. At which level stimuli should be familiar to infants for them to be able to take advantage of familiarity and to generalize rules is an empirical question. Familiarity may refer to specific items and whether they have been heard or seen before (e.g., specific phonemes, syllables, tones, shapes, or pictures). It may also refer to a more linguistic familiarity; for example, with respect to picture stimuli, it can refer to whether the participant has a lexical representation for the picture.

Saffran et al. (2007) claimed that failures and asymmetries in generalization in the visual domain (e.g., S. P. Johnson et al., 2009) may have arisen because those experiments had used shapes as stimuli, which were unfamiliar to infants of that age. The authors argued that if infants are familiar with the visual stimuli making up the Marcus-type rules, they should be able to learn them. To that end, they showed that seven-month-old infants were indeed able to generalize an XYX or XYY rule when they were familiarized with stimuli consisting of images of different dog or cat breeds (Ferguson & Waxman, 2015 showed similar results with infants between three and four months of age).

Saffran et al. (2007) argue that the success of the infants in this study can be attributed to the fact that infants are familiar with dogs and cats at this young age, and that pictures thereof are therefore categorizable for them. Learning with shapes is difficult because young infants have typically not yet been exposed to shapes as such. According to Saffran et al. it is the familiarity of speech that facilitates rule learning with speech stimuli, not the fact that speech is in the linguistic domain.

Yet the question remains, at what level were the dogs and cats in Saffran et al. familiar? Since infants are indeed often familiar with dogs and cats from very early on, they might even have an early lexical label for them (Bergelson & Swingley, 2012). If this is the case, it might mean that the visual rule learning task in Saffran et al. was more linguistic than intended. We thus hypothesized that if the reason for success in the Saffran et al. study was the fact that the infants' lexical knowledge facilitated categorization of the stimuli, rule learning may still be specific to language.

Chapters 2 and 3 both address issues related to differences in learning abilities as a result of the familiarity of the stimuli. In chapter 2, where we focus on rule learning in Marcus grammars from auditory stimuli, we use both speech (familiar) and birdsong (unfamiliar) stimuli. In chapter 3 we attempt to disentangle familiarity at the item level, i.e., stimuli "being familiar" from familiarity at the lexical level, i.e., stimuli "being lexically specified," by comparing the performance of 12- to 14-month-olds on a visual rule learning task in three conditions. In one condition the visual training stimuli consisted of images of familiar objects that were also typically lexically familiar to 12- to 14-month-olds; in another condition the images were of familiar objects that were typically not lexically familiar to 12- to14-month-olds; in the third condition images of nonsense objects were used, which were thus unfamiliar at both levels.

1.5 Ecological validity

It is striking that humans (both infant and adult) do not readily learn and generalize the Marcus-type grammars under all circumstances, while they are able to passively learn more complex artificial grammars (e.g., adults can learn finite state grammars such as BROCANTO which mimic the syntactic rules and word class distributions of natural languages, Fitch & Friederici, 2012), and infants learn the grammars of their native language. Another important factor that plays a role in the learnability of an artificial grammar like the Marcus-type one thus seems to be its ecological validity, referring to the naturalness of the input and the learning environment provided to the participant. Ecological validity may be addressed on multiple levels. Here, specifically, we look at the amount of variability present in the input, the type of learning task provided, and finally, the complexity and naturalness of the grammar itself.

1.5.1 Variability

Natural language contains a great deal of variability at all levels of input. The role of variability in language learning has therefore been studied in experimental settings too. Rost and McMurray (2009), for example, showed that 14-month-olds' learning of minimal pair words was improved if the words were produced by multiple speakers. Gómez (2002) showed that both adults and infants learned a non-adjacent dependency rule more reliably when the familiarization stimuli contained a large number of different units in-between the two units constituting the non-adjacent dependency. Gerken (2006) found that while nine month olds who were presented with a subset of four different triads from the Marcus et al. (1999) stimuli could generalize, infants who were presented with either XX*di* or X*di*X triads did not generalize to novel stimuli. In this case, they learned the specific rule of XX*di* or X*di*X. Having a single item in the Y position (i.e., no variability) thus seems to impede the creation of a Y variable, necessary for rule generalization.

In this thesis, the role of variability in the learning of Marcus-type rules is addressed in two ways, namely by paying attention to the phonological variability of the training stimuli (in chapter 2), and by varying the amount of different familiarization triads (in chapter 5).

Regarding phonological variability, the syllables used in the triads in Marcus and colleagues' original study, as well as in later extensions (e.g., Gerken, 2006) showed very little phonological variability: the familiarization stimuli contained only two different vowels, both of which were front vowels. Three of the four familiarization consonants had a coronal place of articulation, while all test consonants were stops. Four of the 12 triads had identical vowels when syllables were combined (*wilili, wididi, lewewe, dewewe*). One triad had identical consonants when syllables were combined (*wiwewe*). While infants in this experiment did show evidence of learning and generalization, the stimuli are hardly representative of the type of variety of phonemes and syllables found in natural language; we tried to remedy this in our extensions of the Marcus et al. paradigm in chapter 2.

Regarding amount of variability at the triad level, Gerken (2010) showed that when infants were first familiarized with the specific XX*di* rule and subsequently familiarized with five extra trials, two of which followed the XX*di* sub-rule, and three of which contained different syllables in the Y position, they were able to generalize to novel test stimuli following an XXY rule. The presence of some variability in the Y position thus allowed infants to switch from a local to a broad (algebraic) rule. Notably, only three relevant examples were sufficient for allowing generalization to occur. Gerken's later work even showed that under specific circumstances, generalization from a single repeated triad is possible (Gerken et al., 2015). In chapter 5, we try to get more insight into the necessary amount of variability for learning, by presenting participants training stimuli consisting of either three repetitions of 15 different triads, or 15 repetitions of three different triads.

1.5.2 Task and instruction

The last two chapters of the thesis, chapter 5 and chapter 6, discuss experiments performed with adult participants.

There have been few examples investigating Marcus-type rule learning in adults. This may be because we can successfully test adults with much more complex input, as has been shown from a wealth of experiments (e.g., Fitch & Hauser, 2004; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Reber, Kassin, Lewis, & Cantor, 1980), which are summarized in more detail in the introductions of chapters 5 and 6. While Marcus-type rule learning has been shown to be successful in adults (Christiansen, Conway, & Curtin, 2000; Endress et al., 2007; Sun, Hoshi-Shiba, Abla, & Okanoya, 2012; Chen et al., 2015), adult participants in this task seem to be highly susceptible to the way the task is presented to them. In Christiansen et al. (2000)'s successful replication, adults needed to be instructed that they were participating in a pattern recognition task; they were unsuccessful otherwise (Christiansen, personal communication). There is reason to believe that for adults, rule learning and generalization may hinge on how they are instructed about the type of task they have to perform, and on whether they receive feedback or not during the familiarization or test phase of the experiment, particularly when it comes to simple grammars. While more complex grammars have been found to be more easily learned implicitly (Reber, 1976), simpler grammars, in which a variety of different strategies can be applied, may require the participants' attention to be directed to relevant features more explicitly. In natural language acquisition (e.g., in second language learning tasks), we also know that adults typically struggle with learning grammar in an implicit manner. Chapter 5 of this thesis systematically explores the effects of both task instruction and the type of task on the rule learning abilities of adults.

1.5.3 Grammar complexity

Properties of the grammar itself form another aspect that can influence learning. As mentioned above, adults are able to learn more complex grammars than Marcus-style grammars. Therefore, at first glance, it is surprising that adults might only be able to learn these simple grammars under specific task and instruction conditions. On the other hand, Marcus-style algebraic rules, and particularly the XYX rule, hardly reflect naturally-occurring phenomena in natural languages (see also the general discussion in chapter 2).

Lindenmayer grammars, or L-systems (Lindenmayer, 1968) have been used to investigate adults' ability to generalize rules (Saddy, 2009; Shirley, 2014). These grammars were originally developed to describe the cell growth of algae (Lindenmayer, 1968; Lindenmayer & Rozenberg, 1972) and have been applied to other plant structures (Prusinkiewicz & Lindenmayer, 1990; Samal, Peterson, & Holliday, 1994). Their recursive rewrite rules make them suitable for studying language processing as well, as their output is too complex to learn by memorization of surface forms, and their underlying hierarchical rules parallel structures found in human language. For this reason, Saddy (2009) and Shirley (2014) began to explore these grammars with promising results. While adults in Shirley's work could discriminate the L-systems from non-L-system foils during testing, it was not clear on what basis this discrimination was taking place; the L-systems in the experiments were composed of two syllables, *bi* and *ba*, and according to the authors the surface manifestations of these grammars produced a rhythmic pattern that participants may have used to discriminate the L-system grammar from a foil grammar. In chapter 6, we extend on this work and attempt to disentangle the roles of task (continuing in the line of chapter 5), grammar complexity, and rhythm. Using rhythmic sounds instead of syllables, we touch on the issue of domain generality of rule learning (brought up in chapters 2 and 3 with infants), and we explore whether Lindenmayer grammars, being closer to natural language grammars, could be a promising alternative artificial grammar to the Marcus-style grammar for future rule learning studies.

1.6 Summary

With few exceptions (Dawson & Gerken, 2009; Johnson et al., 2009), empirical work on rule learning has been geared towards showing what infants can do, rather than providing a detailed pattern of successes and failures across ages. (Frank & Tenenbaum, 2011)

This thesis begins by tackling the problem identified in the above quotation, exploring in depth failures, as well as a few successes, in rule learning in several age groups and domains. In the end, I will bring together the findings of the five experimental chapters outlined above and discuss findings relevant to the theory of rule learning as well as observations related to methodological considerations.