Simple rule learning is not simple:
Studies on infant and adult pattern perception and production

Andreea Geambaşu

Summary

The work presented in this thesis is part of an interdisciplinary project aimed at understanding the cognitive roots of language acquisition. The goal of this work was to understand the underpinnings of grammatical rule learning by humans: whether our ability to learn our complex human languages is purely specific to us, whether the building blocks of learning language is a general ability that can be applied in other domains (e.g., the auditory domain in general, the visual domain), and how the way a learning task is presented to participants may affect their ability to learn from the input they receive. In the larger project, our collaborators aimed to understand how human abilities might compare to those of other species, namely zebra finches and budgerigars, and to create a comprehensive computational model that can help explain the principles underlying learning of segments of information and the rules that dictate their structure.

This thesis is a collection of papers tackling rule learning from various perspectives, in both infants and adults. Using the technique of creating highly controlled artificial "languages" (referred to in general as an Artificial Grammar Learning paradigm), we were able to gain some insights into learning processes of young infants within their first year of life and of adults, mostly of university age.

Following a general introduction, chapter 2 reports on an attempt to replicate an often-cited set of experiments from Marcus and colleagues (1999) showing that infants as young as seven months old can quickly learn simple patterns following either an XYX, XXY, or XYY rule and recognize those rules when they are presented with speech sounds that were not heard in their learning phase of the experiment. This study has been cited time and again as providing evidence that infants are able to generalize, or transfer, a learned rule to novel information. Yet our large-scale attempt at extending on, and eventually replicating closely, the Marcus et al. (1999) study, was unsuccessful in replicating the original results. Instead we found strong and consistent evidence for infants having a preference for patterns including immediate repetition (XXY or XYY). It didn't make a difference if the sounds that infants heard were speech or non-speech (in this case, short segments from zebra finch songs akin to syllables in our speech).

Another study, reported in chapter 3, focused on whether the ability to learn similar patterns in the visual domain would be improved if infants knew the name for the

objects forming the pattern. A set of experiments was conducted with infants between 12 and 14 months old who are just beginning to speak and who already know the meanings of some common words. Infants were exposed to XYX or XYY rules formed by sequences of three pictures of either objects for which they knew the names, objects that were familiar but for which they did not know the name, or nonsense objects they could not have seen before or have a name for. Contrary to expectations, there was no improved performance as a result of word or object familiarity in terms of rule learning ability. Infants did, however, show more attention for pictures of objects of which they knew the name during the learning phase. Some evidence for a repetition bias was also found here, but it must be further investigated, as the number of participants was small here.

A third study on infants, discussed in chapter 4, dealt with their productions and the patterns found therein. Over 60,000 utterances of children from eight different languages in their first two years of life were analyzed to better understand the types of syllable sequences infants produce during the time we probed their rule learning abilities in perception studies like the ones mentioned above. The utterances revealed that, contrary to previous theories, infant babbling does not follow a specific timeline, in which babies first produce mostly utterances containing repetitions like *bababa* followed by a period in which they produce mostly variegated, non-repetitive utterances like *badigo*. Instead, infants are producing the latter type of utterances most frequently, and from the beginning, along with the former type (called reduplicated utterances). While this is an important finding in its own right, it also provides a relevant reference point for understanding how infants may be processing the patterns we ask them to learn in the Artificial Grammar Learning studies detailed above. Previous studies have shown that infants first repeat novel syllables in their repertoire, before using them in variegated sequences, making repetition an important learning tool for babies. As such, it follows that repeated utterances are also the ones that attract attention in the AGL experiments described above.

In the last two chapters of this thesis, chapters 5 and 6, attention is turned to adult participants and their ability to learn rules. In a simple rule learning task using the same speech stimuli as in the first infant rule learning experiments, adult participants were able to generalize the grammar they were exposed to, but only under specific conditions. The way the task was presented to them determined their success rate. They could not generalize the simple rules when instructions did not specifically tell them to seek a pattern in the input. They also benefitted from participating in an active learning task that involved feedback, rather than a passive learning task. Finally, adult participants were able to generalize from their exposure input to novel test stimuli if the test stimuli were composed of novel items only, instead of a mix of novel and familiar exposure items; this type of testing apparently allowed them to focus on one level of processing -- the structural level. The results of these experiments showed that, as is the case for infants, simple rule learning is not simple for adults either. A final exploratory experiment made use of a far more complex grammar, a Lindenmayer grammar, involving hierarchical structures and therefore more akin to natural languages than the simple grammars detailed above. In two

types of testing paradigms, participants showed evidence of sensitivity to this type of grammar instantiated with different types of drum sounds, indicating that humans do have complex rule learning abilities, in the general auditory domain as well, but that the experimental stimuli and task must be presented in such a way that these abilities can be manifested.