Individual minds create universal patterns

Exploring word order universals using

an artificial language learning experiment

B008345

MSc Evolution of Language and Cognition

The University of Edinburgh

2011

Voor Opa

Ik hoop dat ook deze scriptie bij je in de boekenkast staat, ergens.

Contents

Acknowledgments

There are a number of people who have been of great help in the process of writing this dissertation. I would shortly like to thank them here. First of all, I am very grateful to have had Kenny Smith as my supervisor. He has helped me enormously by generating many ideas (and solutions!) for my dissertation and especially my experiment. Also, his reassuring attitude ('you have loads of time!') when I was silently panicking over my time schedule was always more than welcome.

Various other people have helped me with a lot of the hands-on problems that needed to be solved in the making of my experiment. Enormous recognition goes out to Eddie Dubourg, who has probably answered about a hundred E-Prime questions tirelessly. Ziggy Campbell and Lily were kind enough to help record my sentences, despite the torture-like nature of the procedure. I am also very grateful to Elizabeth Wonnacott for taking time during her busy visit to Edinburgh to give me advice on my experiment. Finally, Matthijs served as the director and camera man for the videos and was there for lots of pep talks along the way. An enormous 'thank you' to all of you!

Abstract

In recent years, the interest of language universals for researchers of language evolution has been pointed out in various publications. Word order universals are particularly robust examples, and the present dissertation focuses on the relation between the order of the verb and the object and the adposition and the noun. In the languages of the world, there is a preference for consistent ordering of the heads in these constituents. An artificial language learning experiment was conducted, in which participants learned one of four different miniature languages, which represented each of the possible order combinations. The results showed a combined influence of an ordering bias and the native language, showing that the head-order bias involved is strong enough to overcome native language influence. Additional interesting findings on the driving force of the verb phrase order towards other word orders are also discussed. The present study lends support to the hypothesis that language universals are the manifestation of biases on the individual level.

1. Introduction

Language universals and especially word order universals have received interest in linguistics for many years. Typologists have identified numerous patterns across the languages of the world which suggest that language structure is not simply random. Particularly word order universals have received much attention as they seem very robust (Greenberg, 1963; Hawkins, 1983; Dryer, 1992). Typological explanations for these word order universals have mainly appealed to functional biases within the languages themselves (Hawkins, 1983; Dryer, 1992).

More recently, researchers have also pointed out the interest of universals from a language evolution perspective (Christiansen, 1994; 2000; Kirby, 1998; 1999). The patterns that emerge in the thousands of languages we have today may well hold important clues to the way humans learn and process language. It has been argued by various researchers that artificial language learning experiments are excellent tools to help us gain a better understanding of potential biases the lie at the basis of language universals (Christiansen, 2000; Tily & Jaeger, to appear). In such experiments, artificial languages are created to isolate particular language features. This way, it can be tested whether acquisition results on these features are similar to the patterns found in the languages of the world. Indeed, various artificial language learning experiments have already successfully replicated universal patterns in various domains (Culbertson & Smolensky, forthcoming; Tily, Frank & Jaeger, 2011; St. Clair, Monaghan & Ramscar, 2009; Christiansen, 2000; Cook, 1988).

This dissertation investigates a potential bias behind the implicational relation between the order of the verb phrase and the adpositional phrase. This universal is a particularly robust example of the series of well-known word order universals that are associated with head-ordering. The universal has received attention in previous studies, which have led to mixed results (i.e. Cook, 1988; Christiansen, 2000). The present study contains an artificial language learning experiment that uses linguistic stimuli and a well-controlled experimental paradigm. Furthermore, by looking at acquisition of all four possible language types involving this order, a detailed comparison of each order combination is made.

The dissertation is structured as follows. In Chapter 2, I will provide an overview of the literature on this topic. I will discuss language universals in general and the specific characteristics of the universal that is in focus. Then I will turn to the various explanations that have been proposed for the existence of universals, particularly word order universals. Recently raised doubts about the validity of language universals will also be considered. Then, I will describe a series of previous artificial language learning studies that are of

relevance for the present investigation. Finally, a short description of the experiment and predictions for the outcome are provided. In Chapter 3, I will describe the design and execution of the experiment in more detail. Next, in Chapter 4, the results of the experiment will be presented. The results will be interpreted and compared to other studies in Chapter 5. Finally, summarizing conclusions will be drawn in Chapter 6.

2. Theoretical framework

2.1. What are language universals?

Language universals have received interest in the field of linguistics for quite some time, and two main approaches are usually identified: generativism and typology (see for example Evans & Levinson, 2009; Comrie, 1981). Generativist accounts have formulated language universals as an innate set of rules, which humans use to learn language (called Universal Grammar, or UG). Language universals are said to arise from this rule system (i.e. Chomsky, 1986). However, as Evans and Levinson (2009) stress, these rules are not necessarily based on extensive research about common characteristics of languages. Within the typological tradition, statements on language universals are distilled from substantial empirical observations of patterns in the languages of the world. Therefore, I will follow the definitions of the typological tradition, as empirical observations are the only valid way of identifying language universals.

As the name suggests, language universals are patterns that occur across all or most languages. The universals that typologists are interested in can be divided into four groups by means of two dimensions. They are either absolute or statistical, and either unconditional or implicational. Absolute universals are exceptionless, whereas statistical universals reflect tendencies across languages. Unconditional universals do not refer to the presence of any other features in the language, whereas implicational universals do. In implicational universals, the presence of one feature is thus associated with the presence of another feature (Comrie, 1981). Universals can apply to all language domains, such as phonology, semantics and syntax. Some of the strongest findings have been done with regards to implicational word order universals. Already in 1963, Greenberg identified a number of word order universals from a small sample of thirty languages. Many of these have been confirmed with larger language samples (e.g. Hawkins, 1983; Dryer, 1992).

2.1.1. The VP/PP word-order universal

The present investigation will be concerned with a very robust example of a word order universal, namely the relation between the order of the object and the verb, e.g. the word order in the Verb Phrase (VP), and the Adpositional Phrase (PP, for Prepositional/ Postpositional Phrase). I will refer to this as the VP/PP universal. This universal is of great interest because it is so robust and because it has been addressed by other artificial language learning experiments, with different results. These will be discussed further below.

The implicational VP/PP universal involves a relationship between the ordering of object and verb (henceforth O and V) and languages having either prepositions or postpositions. Table 2.1 shows their distribution in languages across the world, according to the World Atlas of Language Structures database (WALS online, 2011a).

	Prepositions (512)	Postpositions (577)
Verb-Object (705)	456	42
Object-Verb (713)	14	472

Table 2.1. Distribution of VO/OV order and PP order across languages. The numbers in brackets represent the number of languages in the entire database that have this specific feature (adapted from WALS online, 2011a, which combines Dryer, 2011a; 2011b).

Table 2.1 shows that while VO/OV order and prepositions/postpositions are fairly evenly distributed across the languages of the world, their combinations are not. A clear preference exists for OV + postpositions and VO + prepositions. This pattern was identified by Dryer (1992), in his sample of over 600 languages.[1] Moreover, for a number of languages in the uncommon categories, it has been found that they are actually in the process of a word order change. In these cases, the order of the object and the verb has been changed, but the adposition order is still in the old state, which causes the unusual combination (Dryer, 2011d).

Of course, Table 2.1 oversimplifies the crosslinguistic picture. Some languages in fact do not have dominant word orders within the VP or the PP, and some languages have inpositions instead of pre- or postpositions, or lack adpositions altogether (Dryer, 2011a; 2011b). This is also the reason why the numbers in the table do not add up exactly (aside from the fact that not all languages in the database have been fully described and therefore, in some cases one of the two orders is not known). However, the table above represents the vast majority of languages for which both data on VO/OV and adpositions are available (over 85%). The pattern is thus indeed very strong and wide-spread.

The VP/PP universal is said to be bidirectional, which means that both word orders will imply each other. In the case of the VP/PP universal, this means that having Verb-Object order in the VP will imply that the language will most likely have prepositions, and having prepositions implies that the language also most likely has VO order (Dryer, 2011d). However, as Dryer (2011d) observes: '*It appears to be the case that if a language changes the order of one of these two features, it will almost always be the order of object and verb that*

[1] Dryer thus provides a generalisation of Greenberg's observations on dominant order of the Subject (S), the Object and the Verb and adpositions. In his 30-language sample, Greenberg (1963) found that VSO languages are always prepositional and SOV languages are usually postpositional.

changes first.' This would seem to indicate that while the implication between the VP/PP order is bidirectional, the VP has a leading role. However, no strong claims about either order having a stronger influence on the other have been made in the literature. The two are only said to correlate (e.g. Dryer, 1992; 2011d).

The VP/PP universal is part of a series of word order universals which are all correlated. Some examples of these word orders are the order of the noun and the genitive, the noun and the relative clause and the verb and the PP (Dryer, 1992). It has been argued that these universals occur because there is a preference in languages for consistent head-ordering. This means that the heads of clauses or phrases, such as the verb in the VP or the adposition in the PP, are always in the same position. Usually, languages will thus be either head-initial or head-final, grouping all heads at the same side of phrases and clauses (Tallerman, 2005). A language which has VO and prepositions is thus head-initial, whereas a language with OV and postpositions is head-final.

2.2. Explanations for language universals and their link to language evolution

Even though absolute language universals have not been (and may never be) established (Evans & Levinson, 2009), the robust tendencies we find in languages across the globe have been argued to provide a window on the mechanisms humans use to acquire and process language. Various views on this matter can be found in the literature.

Within the generativist tradition that was shortly mentioned earlier the universal properties of language have been seen as the manifestation of the abstract formal constraints that constitute Universal Grammar. This Universal Grammar describes the set of all possible human languages, and is innate. Children are hypothesized to use this device to learn their mother tongues (Chomsky, 1986). Universal patterns in the languages of the world are thus considered to be the result of these linguistic biases that are assumed to be present in the brain[2]. While this approach has been very influential, it has also been surrounded by a lot of discussion and criticism. Evans and Levinson (2009) recently argued that the generative approach does not use substantial typological evidence to back up their claims, and that a lot more diversity can be found than generativists subsume in UG.

Another argument against an innate linguistic rule system has been put forward by Christiansen and Chater (2008). They argue that no language-specific rule system such as

---

[2] Although interestingly, Chomsky (1980) has claimed that language universals can be derived from a single language, as any characteristics in any language have to originate in the abstract underlying structure that is Universal Grammar.

Universal Grammar could have arisen from biological adaptation to language usage. Their main reason for this is that languages change so fast, that the much slower process of genetic change would not allow for adaptations to accommodate these systems. Christiansen and Ellefson (2002) add that the selective pressures working on languages to adapt to humans are much greater than those working on humans to adapt to language (as humans do not need language to survive, while languages definitely need humans to survive). Languages are thus seen as organisms that are subject to natural selection, and this selection is assumed to be exerted by the human brain.

Instead, then, Christiansen and Chater (2008) argue that languages must have adapted to constraints that are not language-specific: domain-general biases. Various types of cognitive constraints to which languages have had to adapt have been proposed, and many of these have been associated with language universals. In the discussion of empirical studies on language universals we will come across a few of these. A third possibility, besides domain-general or domain-specific biases, was suggested by Kirby (1998, 1999). He proposed that while language may initially have depended on domain-general biases, domain-specific biases may have emerged along the way, to adapt to the demands of language acquisition.

The theories discussed so far share the common property that they suppose some sort of bias in our brains is causing universals to emerge, whether linguistic, non-linguistic or both. In this dissertation I would like to favour the views of Christiansen and Chater (2008). Indeed, it is difficult to see how language may have emerged if it would need an elaborate domain-specific brain structure to be processed by humans.

That said, an important note that needs to be made here is that it is very difficult to tell from artificial language learning experiments which of these accounts is correct. In the literature, strong claims have been made on whether the supposed biases are linguistic (e.g. Chomsky, 1986) or domain-general (e.g. Christiansen, 2000; Christiansen & Ellefson, 2002). Especially artificial language learning studies have made strong claims about constraints being non-linguistic (i.e. Christiansen, 2000). In fact, it is very hard to prove using artificial language learning experiments which of the two is correct. This will be further illustrated below, in the section on artificial language learning experiments. In the present investigation, the focus will therefore be on whether a bias for a specific word order universal can be found or not, whether linguistic or non-linguistic.[3]

---

[3] Interestingly, various modelling studies have shown that very simple connectionist models reproduce universal patterns when they are required to learn natural and unnatural languages (performing better on the natural types). This was shown for head-ordering (Christiansen & Devlin, 1997) and the relation between dominant word order

A different type of theory on the emergence of language universals does not include individual biases, but proposes that language universals are created by historical processes of language change. In particular, it has been argued that the process of grammaticalization may have a prominent role in the emergence of word order universals. Through grammaticalization, content words can become function words. In this process, the words lose meaning and often change form. It has been argued that the common correlations between word orders are linked through history by grammaticalization, as word functions change but their position remains the same. Bybee (1988) describes how adpositions can develop from the head-noun in a genitive construction. For example, 'by the side of-the-house' would develop into 'beside the house' (noun-genitive and prepositions). This would thus automatically match up the order of the two structures, with the two heads being on the same side. Dryer (2011d) posits a similar argument, saying that verbs may also develop into adpositions. It is clear that in an artificial language learning experiment, such processes do not take place, as they take many generations of transmission. In the present experiment this theory will therefore not be explicitly addressed, however, if no individual bias can be found using an artificial language learning experiment, that may be a sign that diachronic processes are indeed more important. Finally, it is also possible that these diachronic processes occur alongside and maybe reinforce the surface effects of an innate bias.

2.2.1. Do language universals really exist?

In recent publications, the existence of language universals and thereby some theories on the causes of language universals have been disputed. Evans & Levinson (2009) argue that in any case, there is no such thing as an absolute universal, as so far all absolute universals that have been posited have met with counter examples. However, they do maintain that strong tendencies that are found across languages are likely to point to *'a cognitive, communicative or system-internal bias towards particular solutions evolving'* (p. 439).

A more critical stand is taken in another very recent article by Dunn et al. (2011). They use phylogenetic methods to investigate how universal language universals really are. They investigated whether eight different word order features could be found to have evolved in the same way across four different language families (Austronesian, Bantu, Indo-European and Uto-Aztecan). It has been difficult for studies on universals to control for the ever-present

and case marking (Lupyan & Christiansen, 2002). These connectionist models did not have any linguistic biases, only a general bias for sequential processing. The fact that such models can replicate universal patterns seems to indicate that no specific linguistic bias is needed.

relations between languages. Dunn et al. intend to control for this by investigating correlations in the evolution of specific word-order combinations. By means of powerful statistical methods, they studied the evolution of these traits across language trees. To do this, they looked at changes in word orders through time, and compared them between families. They claim that their results show that no word-order universals hold across the four language families that were investigated. Furthermore, they show that even if some universals are encountered in more than one language family, these have not evolved towards their current states in the same way. They therefore argue that word order universals are in fact lineage-specific and cannot be caused by any cognitive constraints.

However, the methodology and assumptions used in the article by Dunn et al. have met with criticism. Firstly, the limited number of similarities that could be found between the language families may be due to limitations of the data that compromise the analysis. As Croft et al. (to appear) point out, the arbitrary threshold to balance out Type I and Type II errors that Dunn et al. used has clearly allowed for a large number of Type II errors to occur. These false negatives are most obvious in the data for the Bantu language families, where 18 out of 28 of the features are in fact stable, i.e. no word order changes have occurred for these features. This makes it impossible for the Bayesian analysis to detect word order correlations in this language family. Croft et al. suggest that the Type II error rate is sufficiently large to question the validity of the results. Croft et al. point out a few more problems with the assumptions in the study, such as the fact that the Bayesian traits analysis always assumes phylogenies to be independent. For this study, this means that the four language families are assumed to have no common ancestor, i.e. this is in line with a polygenesis view of the languages of the world. However, Croft et al. point out that the common belief among linguists is that all languages have one common origin. Finally, Dunn et al. did not include the effects of language contact in their analysis. Croft et al. argue that this is a very important factor in language change and therefore, overlooking language contact could potentially influence the results to a large extent, especially because changes in word order are used to track the evolution of the orders.

Altogether then, while the Dunn et al. paper has provided a first step in a potentially important approach to investigating language universals, the assumptions made in the study (some inevitable due to the form of the model) have various setbacks. Claiming that language universals and cognitive biases leading to language universals do not exist therefore seems unwarranted at this time.

The article by Dunn et al. has also again led researchers to advocate the importance of artificial language learning studies. Summarizing a series of artificial language learning studies on language universals, Tily and Jaeger (to appear) argue that artificial language learning experiment are an important complimentary research method to phylogenetic analyses. This is because phylogenetic analyses are seriously hampered by the sometimes sparse typological data. As will be discussed below, various artificial language learning experiments have been able to replicate language universals in the lab, which seems hard to explain if these universals were in fact lineage-specific. Below I will discuss a series of artificial language learning experiments that have considered word order universals. This will provide a feel for the way these studies are designed and what their results have been, and will also further stress the importance of the current study.

2.3. Previous empirical studies on word order universals

Often in an attempt to investigate which cognitive biases affect language structures around the world, various studies involving specific language universals have been conducted. The studies I will discuss here all focused on word order.

In a recent study, Culbertson and Smolensky (forthcoming) used an artificial language learning experiment to investigate whether learning biases are involved in the ordering of adjectives and nouns and numerals and nouns. Looking at the universal distribution of these features, Culbertson and Smolensky hypothesize that there is a cross-linguistic preference for harmonic patterns, in which the nouns are in the same position in both phrases. Also, there is a difference between the two patterns when the position of the noun is inconsistent. Noun-Adjective + Numeral-Noun are more common in the world and therefore possibly unmarked, as opposed to the inverse combination, Adjective-Noun + Noun-Numeral. The universal ordering patterns involving these two phrase types can thus not be attributed to head-ordering preferences.

Participants were trained on artificial languages in four different conditions involving each of the possible ordering combinations. 70% of their input had one specific ordering combination, and 30% had random ordering. In the test phase, participants who learned the harmonic and unmarked combinations turned out to generalize more towards their majority input patterns (e.g. more than 70% of their output contained the majority input pattern). For the marked language, this boosting did not occur.

Culbertson and Smolensky argued that, because these results cannot be attributed solely to head-ordering, a learning bias must be causing this effect. Thus, people generalize

towards these patterns as they are easier to learn, and in this case, these learning biases are in favour of harmonic and unmarked patterns. The authors define 'learning bias' as asymmetric learning of given structures, i.e., in learning, a certain structure can be favoured over another. This definition is thus more like a description of the findings, but no specific claims are made with regards to the domain-specific or domain-general nature of this bias. It is also not discussed why this learning bias would go against the more general head-ordering preferences that are reflected in many word order universals. Either way, the replication of the unusual universal pattern using these small artificial languages shows that people in fact exhibit universal tendencies on the individual level.

Aside from the preference for harmonic and unmarked orders, a slight native-language effect was found as well. In the language which was most disfavoured by the preferences for harmony and unmarkedness, performance on the numerals, which had non-English order, was worse than for adjectives, which had English order. The native language effect was not found across all four languages however, and its influence was thus very small.

In 2000, Christiansen performed an artificial language learning experiment to investigate the sensitivity to head-order consistency. As was mentioned earlier, the VP/PP universal and many other word order universals are associated with a preference for consistent head-ordering languages. Two groups of participants learned two different languages that were either consistent or inconsistent across phrases in terms of the position of the head. The stimuli that were used were strings of letters which were generated using these consistent or inconsistent grammars, where each letter represented a specific constituent. The design of the two languages is shown below in Table 2.2.

Consistent grammar			Inconsistent grammar		
S	$\rightarrow$	NP VP	S	$\rightarrow$	NP VP
NP	$\rightarrow$	(PP) N	NP	$\rightarrow$	(PP) N
PP	$\rightarrow$	NP Postposition	PP	$\rightarrow$	Preposition NP
VP	$\rightarrow$	(PP) (NP) V	VP	$\rightarrow$	(PP) (NP) V
NP	$\rightarrow$	(PossessiveP) N	NP	$\rightarrow$	(PossessiveP) N
PossessiveP	$\rightarrow$	NP Possessive	PossessiveP	$\rightarrow$	Possessive NP

Table 2.2. Consistent and inconsistent grammars in Christiansen (2000) (adapted from p. 47).

Both languages thus had OV order in the VP, but the consistent language has postpositions whereas the inconsistent one has prepositions. The order of the possessive and the noun is also reversed for the inconsistent language.

The participants were first trained on a set of these strings and were then asked to judge a new set of strings on grammaticality. In the consistent condition, participants were able to perform correct grammaticality judgements for 63.0% of the novel sentences, versus 58.3% of the participants in the inconsistent condition (this difference was significant, $p <$ .02). Christiansen attributed this to constraints on sequential learning and processing.

The study thus shows that even when people are presented with non-linguistic stimuli, they perform in line with the evidence from language universals. Christiansen claims that it therefore proves that language universals arise from non-linguistic biases. However, because non-linguistic stimuli are used, the experiment may only be tapping into non-linguistic cognitive skills. This study can thus not be said to prove unambiguously that real language learning is dependent on these non-linguistic constraints as well. This is thus another example of an attempt to formulate the type of bias very strictly even though the evidence does not necessarily allow it. Either way, the fact that this bias is found is a highly interesting result. Christiansen does not report on any native language effects. He does mention that the consistent grammar is head-final, in contrast to English, to avoid native language influence. Whether a possible native language effect would be subordinate to the domain-general bias, as in Culbertson and Smolensky (forthcoming) is thus not assessed.

An early and quite extensive artificial language learning experiment was conducted by Cook (1988). Again, this study focuses on word order and specifically on the ability of learners to extrapolate word order from the VP (e.g. the order of object and verb, the verb being the head) to the PP and the Noun Phrase (the NP, involving an adjective and a noun). I shall focus here on the results from the condition in which learners had to extrapolate from the VP to the PP, as the results for this particular group were not in line with the findings on the implicational relationship between these phrases.

The participants in Cook's experiment were English speaking children who also learned either French or German in school. One group was trained on a set of sentences with SOV order, the other on sentences with VSO order. These orders were chosen to avoid presenting the children with the dominant English word order, SVO. Training consisted of the children's class teachers orally presenting 30 sentences in the target language. They first read out the sentence, followed by the English gloss and finally the English translation. Via multiple choice questions, the children were tested on their acquisition of the presented word order. Then, again via multiple choice questions, children had to read English sentences involving adpositions, and select the correct translation in the target language. Crucially, these

adpositions were not present in the training sample. Both training and testing took place in class.

The extrapolation results for these two languages are summarized below in Table 2.3.

	Prep	Post
V(S)O	35%	63%
(S)OV	7%	90%

Table 2.3. Extrapolation of word order from VP to PP. Adapted from Cook (1988).

The results in Table 2.3 do not correspond to the distribution of VP-PP combinations across the world. In fact, as Cook also points out, there seems to be a preference for postpositions. Cook argues that this could be a strategy the children use for this task, possibly resulting from an attempt to construct "anti-English". Another possible explanation that Cook provides is that the children performed some sort of problem solving, and that they did this remarkably consistently (as most of them came up with the same solution). Unfortunately, it is not possible to identify exactly what specific problem the children collectively saw in the task.

Clearly, then, Cook's results are puzzling and hard to explain. They do not seem to show an underlying bias towards the universal patterns, and it is difficult to find another constraint that may have caused the participants to behave in the way they did. It may be that a number of flaws in the design of the study caused these unexpected results. For example, the method that was used is very different from any natural language learning situation, even from second language acquisition. The children only received glosses and translations for the target languages, and these were presented in a rather unnatural way. Furthermore, the children were tested using written questionnaires with multiple choice questions, possibly leading them to take on the problem in a more analytic way than they would have in normal language learning. Indeed, Cook's proposed explanation of some type of problem-solving seems to point in this direction. In addition, a major problem with this study is the fact that it was carried out in class, which makes for a virtually uncontrolled experimental situation.

Finally, the subject-verb-object orders used in the two languages may have led to the different results, as the order of the subject and the verb are different between the two languages. Cross-linguistically, there is a preference for Subject-Verb order (1194 language in the WALS database) as opposed to Verb-Subject (194 languages in the WALS database; Dryer, 2011e). It is remarkable that for the VSO order, where the verb and the subject are thus not in the preferred order, the 'wrong' adposition is usually selected. It is currently not known why exactly VS is less common (Dryer, 2011e). Nevertheless, the fact that it is uncommon

indicates that it is possibly harder to learn, which may have caused Cook's participants to behave in an unexpected way in this language.

2.4. The present study

The studies reviewed here already show there is large variety in artificial language learning paradigms. What is also clear is that not all evidence is in line with the word order universals we are currently aware of. While Christiansen (2000) included adpositional phrases in his investigation of head-ordering and found that participants behaved in line with typological evidence, Cook's participants (1988) failed to show such a pattern. It seems likely that the methodological flaws in the design of Cook's experiment contributed to this.

The present study will specifically target the VP/PP universal using linguistic stimuli. Hereby, it will be investigated whether learners exhibit the word order universal in isolation (in contrast to Christiansen's study, which investigated overall consistency). It may be that inconsistency is easier to overcome given a small number of word orders, which would indicate that the universals are all tied in together, rather than each having a separate relation with each other order. Furthermore, by using four languages, each possible combination of orders will be tested. This will allow for a more accurate assessment of the strength of a possible bias in comparison to the influence of the native language than in Christiansen's experiment. In the present study, participants will also be required to perform a production task, which will provide us with more insight in the way they use the knowledge they have gained. Finally, it will show whether Cook's unexpected results are replicated using a more standard artificial language learning paradigm with greater experimental control.

2.4.1. Experiment design

Like Culbertson and Smolensky's (forthcoming), the artificial language learning experiment testing the VP/PP universal will involve four languages, together representing all possible order combinations between the two phrases. Two will thus follow the universal pattern, and two will not. The main characteristics of the four languages are presented below in Table 2.4. Further details on the languages are provided in the methodology section.

1	VO	Prepositions	Natural
2	VO	Postpositions	Unnatural
3	OV	Prepositions	Unnatural
4	OV	Postpositions	Natural

Table 2.4. Global structure of the four languages.

In contrast to Culbertson and Smolensky's experiment, in the present experiment participants will receive fully consistent languages. The languages are designed to be difficult enough so that learners cannot acquire them perfectly by means of the limited amount of training they receive. Through a comparison of the test scores, comparative learnability of the languages will be assessed directly.

The stimuli and learning paradigm are similar to those used by Wonnacott, Newport & Tanenhaus (2007). Meanings are represented by short movies, and the language is presented aurally. The movies involve puppet animals performing actions to each other. The movies also include other objects so that reference to location becomes necessary.

The languages are learned in two stages. First, single nouns are learned by means of pictures. Then, the verbs, adpositions and word order are trained by means of videos of two hand puppets which perform actions, sometimes with an inanimate object present near the patient. Comprehension tests, production tests and grammatical judgement tests are assigned to assess performance.

2.4.2. Predictions

In the present dissertation, I would thus like to answer the question *Is the implicational relation between word order in the verb phrase and the adpositional phrase the result of an underlying cognitive bias?*

Based on the universal distribution of the VP/PP orders and findings from other studies, the following predictions can be made with regards to the results.

(a). If there is a bias towards coordination of VP/PP order, then languages 1 and 4 will be easiest to learn and therefore learners of these languages will perform best, leading to the following ranking of the test scores:

1,4 > 2,3

(b). If the native language influences performance alongside the VP/PP universal bias then the ranking of scores will be as follows:

1 > 4 > 2,3

As language 1 is closest to English, being VO and prepositional it should be learned better than language 4, which is the opposite of English on these two language features.

(c). If native language influence is more important than the VP/PP bias, then the ordering would be as follows:

$1 > 2,3 > 4$

As language 4 is most different from English, while languages 2 and 3 share one of the crucial features with English and language 1 shares both of the features with English.

(d). If performance is similar to that in Cook (1988), there should be a preference for anti-English with regards to postpositions, and the ranking of the scores should be as follows:

$2,4 > 1,3$

3. Methodology

3.1. Participants

The participants were undergraduate students of the University of Edinburgh, who were paid for their participation. All participants were native speakers of English and did not consider themselves fluent in any other language.

3.2. Sentences

The basic word orders of the languages were VOS and OVS. These basic word orders were chosen because both are different from the participants' mother tongue, English, which has SVO order. Furthermore, both basic orders have the same order of the verb and the subject, eliminating any possible differences between the languages because of this order. These basic word orders are both highly uncommon in the languages in the word, but the implicational relation with PP order holds for both types, as is shown in Table 3.1.

	Prepositions (512)	Postpositions (577)
VOS (25)	20	0
OVS (11)	3	8

Table 3.1. Distribution of dominant word orders and PP order across languages. The numbers in brackets represent the number of languages in the entire database that have this specific feature (adapted from WALS online, 2011b, which combines Dryer, 2011b; 2011c).

Of course, as so few languages of these types have been described thus far, there is a small risk that in fact, they do not follow the relation between VP and PP order (we may not have enough data to see this). Indeed, a chi-square test of the above table shows however that there is a significant relation between the two orders (Fisher's exact test: $p = .000$). Furthermore, seeing as the relationship is found to hold across all other basic word order types as well (WALS online, 2011b) it is most likely that VOS and OVS are no exceptions to this. The only remark that can be made here is the observation made by Dryer (2011d) that out of the 14 languages that combine OV and prepositions, it is surprising that 3 languages are of the highly uncommon basic word order type OVS. Considering the distribution of basic word orders, this is a fairly high number. However, due to the small numbers involved here, it is likely that it is a coincidence that this distribution occurs in the languages discovered so far. Therefore this observation cannot overrule the advantages that these two orders have for the present study.

The order of the adposition and the noun is also systematically varied across the four languages, with two languages having prepositions and two having postpositions. It is not

possible to avoid the inclusion of a third word order, that of the PP and the noun that is modified by the PP (henceforth N/PP order). The ordering of these two constituents has not been described separately in the literature on word order universals. However, these two constituents do have a head-dependent relation, where the noun is the head and the PP is the dependent that modifies the head. This order is correlated with the order in the VP and the PP. For head-initial languages (VO + prepositions), the order will usually be N-PP, for head-final languages (OV + postpositions), the order will be PP-N (see for example Christiansen, 2000). To avoid additional reinforcement of head order in either of the languages, it was chosen to use both possible N/PP orders in each language and keep the distribution of the order perfectly balanced in all four languages. The four languages thus included the full sentence structures listed in Table 3.2 below. The example sentence 'the lion rams the dog who is behind the box' is glossed in each language to clarify.

1	VOS	Pr-N	N-PP	$V [N_{Patient}[P N_{Inanimate}]] N_{Agent}$ = ram dog behind box lion
			PP-N	$V [[P N_{Inanimate}]N_{Patient}] N_{Agent}$ = ram behind box dog lion
2	VOS	N-Pr	N-PP	$V [N_{Patient}[N_{Inanimate} P]] N_{Agent}$ = ram dog box behind lion
			PP-N	$V [[N_{Inanimate} P]N_{Patient}] N_{Agent}$ = ram box behind dog lion
3	OVS	Po-N	N-PP	$[N_{Patient}[P N_{Inanimate}]] V N_{Agent}$ = dog behind box ram lion
			PP-N	$[[P N_{Inanimate}]N_{Patient}] V N_{Agent}$ = behind box dog ram lion
4	OVS	N-Po	N-PP	$[N_{Patient}[N_{Inanimate} P]] V N_{Agent}$ = dog box behind ram lion
			PP-N	$[[N_{Inanimate} P]N_{Patient}] V N_{Agent}$ = box behind dog ram lion

Table 3.2. Full sentence structures in each of the four languages.

Languages 1 and 4 thus follow the universal, whereas languages 2 and 3 do not. Unfortunately, the randomization of the order of the PP and the noun is not entirely without problems either. In Table 3.3 below the ordering of all components for each language are shown again. The combinations that are left-headed are marked in boldface, right-headed combinations are cursive. As becomes clear, languages one and four are the only ones which include sentences that are fully consistent in their head-ordering (underlined sentence structures). This means that there is a confound between overall consistency and VP/PP consistency. However, the insecurity about the N/PP ordering principle is present in all four languages, so no language is fully consistent.

An alternative solution to this problem was considered, namely using eight instead of four different languages, and thus systematically varying the N/PP order across languages as well. However, that would have taken away the possibility to focus on the relation between

the VP and the PP, and instead would have been more directed at general head-order consistency. Altogether, the balanced distribution of N/PP order is the best possible solution.

1	**VOS**	**Pr-N**	**N-PP**
	VOS	**Pr-N**	*PP-N*
2	**VOS**	*N-Po*	**N-PP**
	VOS	*N-Po*	*PP-N*
3	*OVS*	**Pr-N**	**N-PP**
	OVS	**Pr-N**	*PP-N*
4	*OVS*	*N-Po*	**N-PP**
	OVS	*N-Po*	*PP-N*

Table 3.3. Head-ordering consistency in each of the four languages

3.3. Lexicon

The lexicon contained eight nouns, two verbs and two adpositions, all with meanings that were easily depicted in short, simple videos. The lexicon was kept small in order to keep the vocabulary learnable within a short amount of time. The animals, objects, actions and adpositions were chosen to be maximally different and easily distinguishable in the videos.

The word forms were designed in such a way that each grammatical category had its own phonological form, making the categories easy to distinguish. All words follow the phonological rules of English. Nouns were all disyllabic, with syllable structure CVC and all had different initial sounds. All nouns had word-initial stress. The verbs were monosyllabic, all with the structure CCVC. Adpositions were monosyllabic with VC structure, thus in keeping with the shorter form of function words in natural languages. All meanings and words are listed below.

Nouns: animates: *dog* 'walsid'

    lion 'mernat'

    pig 'poltun'

    zebra 'tifpog'

  inanimates: *basket* 'fadlon'

    box 'kestur'

    pan 'surmip'

    towel 'hindel'

Verbs: *kiss* 'smop'

    ram 'blid'

| Adpositions: | *behind* | 'ul' |
| | *in-front-of* | 'ag' |

3.4. Materials

The pictures and videos were made using hand puppets and realia. All videos were three seconds long, giving participants some time to appreciate the scene before the action was depicted.

The words and sentences were recorded by a speaker with a North American accent and presented aurally. Single words were recorded for the first training phase. For the second and third training phase and the test phase, each word was recorded in each possible position in a sentence with nonsense words consisting of two syllables 'ha'. For example:

tifpog ha-ha ha-ha ha-ha ha-ha
ha-ha ha-ha ha-ha smop ha-ha

The recording in sentences was done to give the final sentences a better flow (as opposed to recording single words). The use of the nonsense syllable 'ha' was intended to avoid unnatural transitions between words due to the spreading of phonemic features. The words were cut out of the sentences and concatenated to produce relatively natural-sounding sentences. This happened online within the programme E-Prime which was used to conduct the experiment.

3.5. Procedure

Participants performed the experiment on their own at a computer. They were taken through the programme automatically and written instructions were provided on screen before each phase.

Phase 1: In this phase the participants acquired the nouns. They were shown pictures of each animate and inanimate noun in random order, while hearing the associated words. Each noun was presented six times to ensure all words were learned well, as this was essential for the acquisition of the sentence structures. The participants were encouraged to repeat the words out loud to themselves to learn the words better. Afterwards, the participants received a short comprehension test in which they had to select the right picture from two options for the word they heard. All words were tested once, so this test consisted of eight items.

Phase 2: In this phase the participants were trained on simple sentences (involving only a verb, an agent and a patient) and complex sentences (involving a verb, agent, patient, adposition and an object). In this round, participants thus learned the verbs and the adpositions, the dominant word order and the order of adposition and noun. The sentences were presented in random order.

The training consisted of short videos (three seconds) that were showing either simple or complex events. In the simple events, one animal performed an action to another animal. In the complex events, an animal performed an action to another animal, whilst an inanimate object was positioned either in front of or behind the second animal (e.g. the patient). After three seconds, the videos remained in their final frame, which was illustrative of the event as a whole. The participants then heard the sentence that described the scene in their language. As in phase 1, participants were encouraged to repeat the sentences to themselves. Pictures 3.1 and 3.2 below show examples of final frames for a simple and a complex event, respectively.

Picture 3.1. Simple event 'the lion rams the zebra'

Picture 3.2. Complex event 'the pig kisses the dog who is in front of the basket'

The participants received 18 simple sentence exposures and 32 complex sentence exposures. In these sets, the occurrence of the two verbs and the two adpositions was balanced (e.g. 9 + 16 exposures to each verb and 16 exposures to each adposition). The occurrence of the animate and the inanimate nouns was also balanced as much as possible. Recall that all four languages have two possible sentences for each complex meaning (because of the balanced distribution of N/PP ordering). Therefore, each complex video occurred twice, once with each possible N/PP ordering. This ensured that there was no difference in meaning between the two structures in the training. The complex training thus consisted of 16 different meanings.

Test phase: Participants performed a comprehension test, a production test and a grammatical judgment test. The comprehension test only included simple sentences. These tested the comprehension of the two different verbs (4 items) and of the order of object and verb (4 items) by means of different distractors. Participants heard a sentence and saw two videos and were required to click on the final frame of the video that corresponded to the sentence they heard. A distractor in the verb condition would involve the same animals in the same roles as in the target meaning, but with a different action. For example, participants would thus hear the sentence 'the dog rams the lion', and see movies of a dog ramming a lion and a dog kissing a lion. For the order of the object and the verb, the distractor would involve the exact same lexical elements, but the agent and the patient would be reversed. I.e., participants would hear the sentence 'the pig kisses the zebra' and they would see videos of the pig kissing the zebra and the zebra kissing the pig. It was chosen not to test complex sentences here as comprehension tests could not provide any additional information with regard to the ordering of the adposition and the noun (as one can still understand the meaning of a sentence without knowing what the order of the adposition and the noun is).

For the production test, participants saw new videos which showed meanings which had not occurred during the training phase. They were required to vocally produce a sentence in their language to describe the video. They were advised that multiple answers could be correct, but that they could only give one. There were 12 novel complex meanings in this test, plus 8 simple items. Productions were recorded using a microphone and saved by the experiment programme for transcription.

The final test was a grammaticality judgment test, which included 24 novel complex items. Half of these items were in the participants' own language, and the other half was in the three other languages (i.e. four sentences per other, 'wrong', language). The test also included 8 simple sentences, again half of which were correct and half were incorrect (i.e. four sentence in each VP order). The ordering of N/PP was balanced for the complex sentences throughout this test, so half of the sentences in each language were PP-N, and the other half was N-PP. Participants heard each sentence twice and were required to click on a picture of a green tick if they thought the sentence was correct, or a red cross if they thought the sentence was incorrect. All sentences expressed different meanings.

After the experiment, participants were asked whether they had ever learned any other languages besides their native tongue, and for how long they had studied these languages.

The presentation of two different sentence types within one training round may seem to be rather demanding. However, pilots showed that training on the verbs and the dominant word order separately using a round with simple sentences only was too easy, resulting in ceiling performance.

Two different meaning sets were made for the experiment. Each meaning set consisted of a different set of videos (and thus sentences) for the sentence training, the comprehension test and the production test, and a different set of sentences for the grammaticality judgments. This was done to ensure there was no bias involved in either the training or the test phases, for example because of some lexical elements possibly being combined more often than others. Participants were randomly assigned to the eight different conditions (i.e. one of the four languages and one of the two meaning sets).

3.6. Coding and analysis

For the vocabulary test, the comprehension test and the grammaticality test, total correct scores were used for statistical analyses. The production task was scored by hand on a series of measures. The total number of correctly produced dominant word orders was calculated. This thus included both simple and complex sentences, which means that the maximum score was 20. The total number of correctly produced adpositional phrases was also counted, which thus only involved the complex sentences, equalling 12 items. The position of the adpositional phrase relative to the animate noun it modified was also considered, and a variable of the percentage of left-positioned PPs out of all produced PPs was calculated. The number of vocabulary errors was counted, and the number of omissions of single words, partial or full PPs and full sentences was counted. Finally, the number of 'PP split' constructions (described in more detail in the next chapter) that some participants used was also counted.

An independent sample T-Test was performed to check whether there was any difference between the two meaning sets that were used. Furthermore, one-way ANOVAs were used to determine the influence of participants' experience learning other languages and a possible effect of sex. Two-way ANOVAs were performed to investigate whether the scores were significantly different for either the dominant word order or the adposition order, and whether there was an interaction between the two. For the scores that showed significant interactions, Tukey post hoc analyses were performed to determine the exact differences between the languages.

4. Results

4.1. Participants

In total, 46 people were tested. For one participant technical problems had occurred during the production test which caused here data not to be recorded. The rest of her data were therefore excluded from the analyses. Eight participants were excluded as outliers, because their production data were very limited. Five of them were excluded because they had omitted seven or more complex sentences, adpositional phrases or parts of the adpositional phrase (either the adposition or the noun). One participant simply failed to remember any of the words and produced word forms of which the majority could not be reliably recognized as the words in the language. Finally, two participants produced 'split' adpositional phrases in over half of their complex sentence productions. They thus inserted the animate noun to be modified between the adposition and the inanimate noun. This structure never occurred in the training data and as these participants used this structure in (almost) all of their sentences, they cannot be said to have learned the language structure adequately, thus making them unsuitable for further analysis.

Altogether, the outliers were thus people who had, for one reason or another, no analysable data for more than half of the complex sentences. This cut-off point was used as such an important lack of unanalysable productions adds too much noise to the data.

Table 4.1 below shows the languages of the participants who were excluded as outliers due to limited production. There are some differences between the languages, but an equal number of participants was excluded from the natural and the unnatural languages.

	Language 1	Language 2	Language 3	Language 4
Number of outliers	1	4	0	3

Table 4.1. Number of outliers excluded per language because of limited production.

In the final analyses, 37 people were thus included. Table 4.2 below shows the details of the participants per language and in total.

	Language 1	Language 2	Language 3	Language 4	Total
Number of participants	10	7	11	9	37
Female participants	4	2	5	3	14
Age	22;2	21;0	20;7	21;7	21;4

Table 4.2. Participants' details per language and in total.

4.2. Pre-analysis tests

Independent samples T-Tests were performed with all test variables to see whether there was any difference between the two meaning sets that were used. No significant difference was found for any of the test variables (largest t-value with associated lowest p-value was: t(35) = -1.636, p = .113), which means that the scores were not influenced by the meaning sets. Therefore the meaning sets were collapsed for the further analyses.

It was found that some of the test variables were correlated with participants' language learning experience (both for the number of years they had learned other languages and the number of other languages they were familiar with). A one-way ANOVA was performed with language as a factor to investigate whether 'language learning experience' was distributed randomly across languages. Both for the number of years participants had studied other languages (F(3,31) = .726, p = .544) and the number of other languages they were familiar with no significant differences were found (F(3,31) = .374, p = .772). The spread of experienced language learners was thus indeed random and therefore it was not included in the analyses as a possible confounding variable.

The sex of the participants was also found to be correlated with some of the test variables, with females outperforming males. However, again a one-way ANOVA with language as the independent variable showed that the spread of males and females was random across the four languages, as there were no significant differences across languages (F(3,33) = .189, p = .903). Therefore it was decided not to include sex as a covariate in further analyses.

4.3. Two-way ANOVAs

Two-way ANOVAs with VP order and PP order of the target language as factors were used to investigate the effects of both PP order and VP order on the various test results. As the results on the vocabulary test, the comprehension test, the simple grammaticality judgments and the VP order in the production task were virtually perfect, they did not show any significant effects of either word order in further analyses. This is clearly illustrated by the charts of the mean scores and standard deviations in Figures 4.1 to 4.4 below. All tables in this section are clustered by VP order, with separate bars for each PP order. The numbers in the bars indicate the languages that they correspond to.
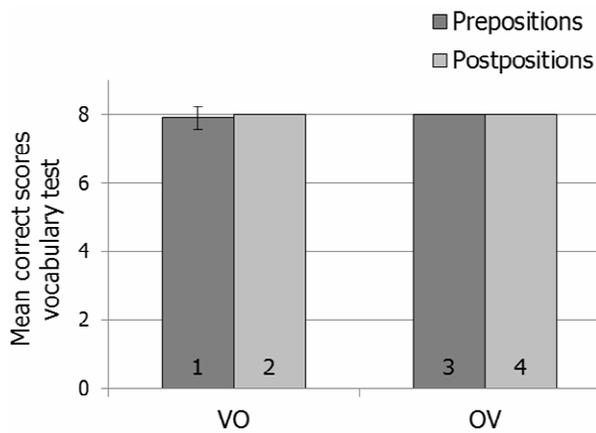
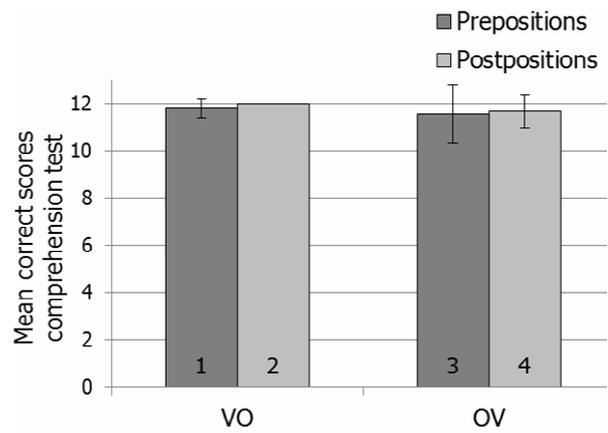Figure 4.1. Means and standard deviations of the correct scores on the vocabulary test (8 items).

Figure 4.2. Means and standard deviations of the correct scores on the comprehension test (12 items).
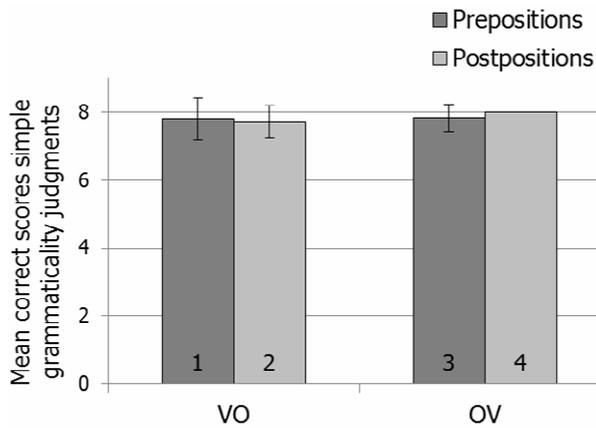

Figure 4.3. Means and standard deviations of the correct scores on the simple grammaticality judgments (8 items).
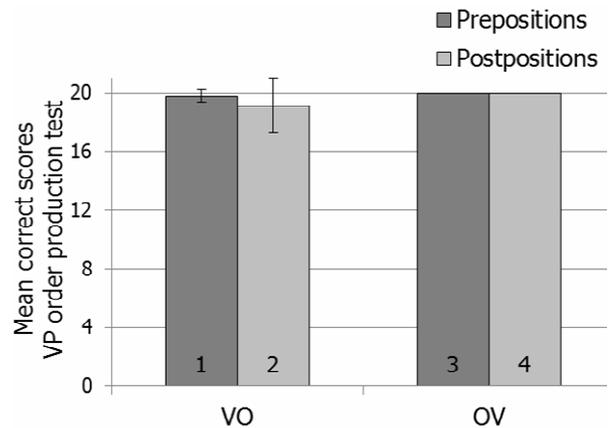

Figure 4.4. Means and standard deviations of the correct scores the VP order in the production test (20 items).

All of these measures are concerned with vocabulary and VP order, but not PP order. Figures 4.1 to 4.4 show that the participants performed virtually at ceiling for all these measures, which indicates that the VP order was well-learned in all languages.

4.3.1. Complex grammaticality judgments

The two-way ANOVA with the total correct score for the complex grammaticality judgments showed a marginally significant interaction between the VP order and the PP order of the different languages ($F(1,33) = 3.986$; $p = .054$). There was no separate significant effect from either the VP order ($F(1,33) = 2.148$; $p = .152$) or the PP order ($F(1,33) = .855$; $p = .362$). The means and standard deviations of the scores are provided in Figure 4.5 below to visualize the interaction.
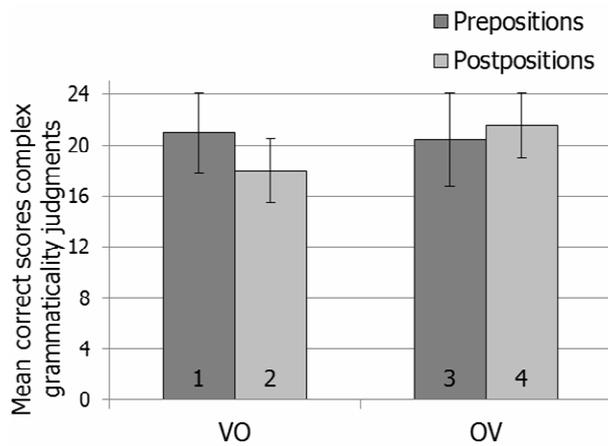
Figure 4.5. Means and standard deviations of correct scores for complex grammaticality judgments (24 items).

Figure 4.5 shows that in general, performance on OV languages and prepositional languages seems to be slightly better, however, this difference was thus not found to be significant. The interaction that is found indicates that the effects of the two factors are dependent on each other. Figure 4.5 seems to show that performance on the natural languages 1 and 4 and the unnatural language 3 is very similar. Performance on language 2 seems to be worse. However, Tukey post-hoc comparisons of the scores per language revealed that there was no significant difference between the scores for any of the languages (lowest $p = .121$). The lowest p-values were observed for the difference between language 2 and the other languages (language 1: $p = .217$, language 3 $p = .366$, language 4 $p = .121$), and p-values for the differences between the other languages were much higher (next-lowest p-value = .856). However, it is clear that no languages reached the α of .05 so at this stage it is not warranted to see this as a genuine pattern. Performance on language 2 is thus in fact not significantly different from the other four languages.

4.3.2. PP order in production task

The two-way ANOVA for correct scores on the PP order in the production task showed a significant effect for adposition order ($F(1,33) = 5.730$; $p = .023$) and a significant interaction between the two factors ($F(1,33) = 7.098$; $p = .012$). The effect of the VP order was non-significant, although fairly close to .05 ($F(1,33) = 3.617$; $p = .066$). The means and standard deviations are again provided below in Figure 4.6.
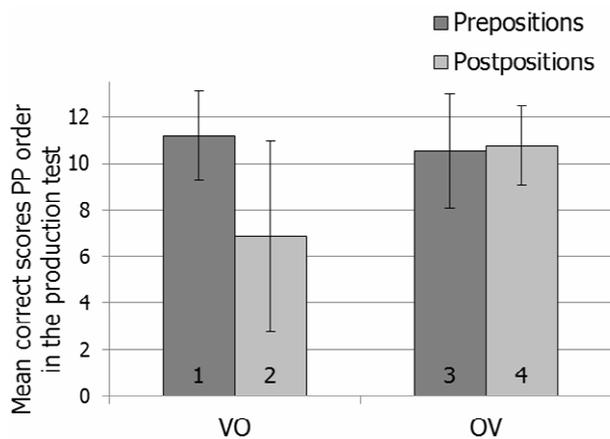
Figure 4.6. Means and standard deviations for correct scores for PP order in the production test (12 items)

The means for the PP order scores in the production test show a picture that is similar to the scores for the complex grammaticality judgments. Again, performance on OV and prepositional languages seems to be better, and this time the difference between prepositions and postpositions is thus indeed significant. The significance of the interaction indicates that the effects of the two orders are again dependent on each other. Again, languages 1, 3 and 4 seem to be very similar, with language 2 showing the worst performance. Tukey post-hoc comparisons showed that in fact, only language 2 is significantly different from language 1 (p = .009), language 3 (p = .027) and language 4 (p = .024). Between the other languages, no significant differences were found (lowest p = .937).

4.3.3. N/PP order

Another variable of interest is the positioning of the adpositional phrase relative to the animate noun, the N/PP order. As was discussed in Chapter 3, in VO and prepositional languages, the PP will mostly occur on the right of the noun it modifies, whereas in OV and postpositional languages, the PP will mostly occur on the left. A two-way ANOVA was performed on a measure of left-positioning of the PP. This variable was calculated by dividing the number of left-placed PPs by the total number of PPs used. It was found that there was a significant effect of VP order ($F(1,33) = 7.442$; p = .010). There was no significant effect of the PP order ($F(1,33) = .178$; p = .675) and no significant interaction between the two orders ($F(1,33) = 2.839$; p = .101). Again mean scores and standard deviations for left-positioning are provided in Figure 4.7. Figure 4.8 shows the average scores per VP type to further clarify the direction of the VP effect.
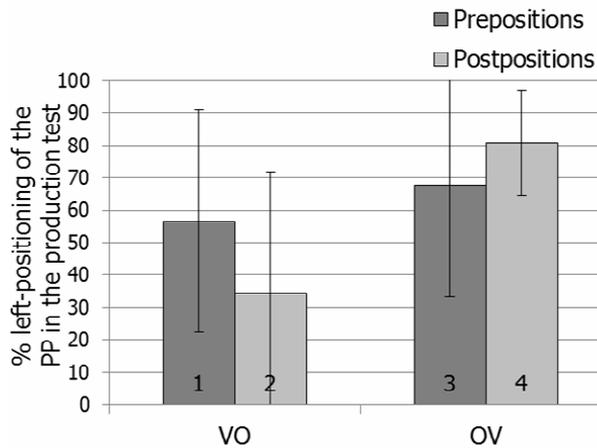
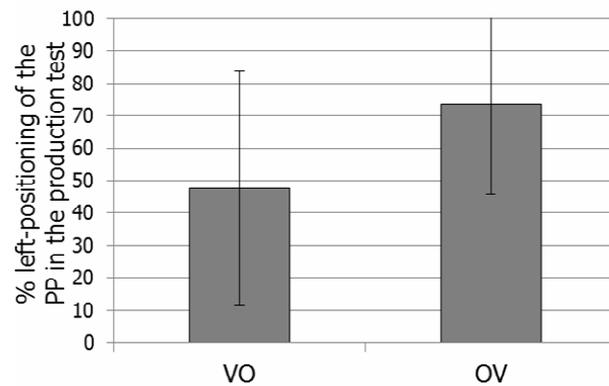Figure 4.7. Mean scores and standard deviations of percentage of left-positioning of the PP by language

Figure 4.8. Means and standard deviations of the percentage of left-positioning of the PP by VP order

Figure 4.8 shows that overall, people who learned a VO language produced close to 50% of their PPs to the left of the head noun. People who learned OV languages placed the adpositional phrases on the left more often. The latter ordering is in line with the OV order, which indicates a steering influence from the VP order, which was thus found to be significant. As mentioned above, no significant effect of the PP order was found.

Seeing as the input distribution of PP position was 50/50, it is interesting to see whether the participants have copied this distribution. One-sample T-Tests with a test value of 50% were used to find out whether the percentage of left-positioned PPs is different from 50% for either VP type or either language. For the VP type, it was found that the percentage for OV languages is significantly different from 50% ($t(19) = 3.793$, $p = .001$), while the score for VO languages does not differ from 50% ($t(16) = -.279$, $p = .784$). This suggests a native language effect which allows people to replicate the trained order better for the VP order that is closest to the one in their native language. When we look at the scores for the individual languages in Figure 4.8, there seems to be a trend where the languages closer to English are also closer to 50% (with language 1 being closest, languages 2 and 3 being further away but at an equal distance, and language 4 being furthest). In fact however, it is found that only the PP positioning in language 4 differs significantly from 50% ($t(8) = 5.751$, $p = .000$). For language 1 ($t(9) = .616$, $p = .553$), language 2 ($t(6) = -1.095$, p = .316), and language 3 ($t(10) = 1.705$, $p = .119$) the PP-positioning did not differ significantly from 50%. This seems to indicate that there was an influence from the native language, as the only language that differed significantly from 50% is the language that is least close to English. Note that there seems to be a trend for lower t-values and higher p-values from language 3 to language 1, which would confirm the pattern of native language influence and differentiation between the influence of

the VP and the PP. However, no claims can of course be made based on such a trend with high p-values.

4.3.4. Vocabulary errors

Another two-way ANOVA was performed to investigate whether there was any difference between the languages in terms of the number of vocabulary errors that were made. These vocabulary errors only include those made during the production task, as a very limited number of errors was made in the initial vocabulary test involving nouns. Figure 4.9 below shows the means and standard deviations. The ANOVA found a significant effect of VP order ($F(1,33) = 4.121$; $p = .05$). Figure 4.9 shows that this means that fewer vocabulary errors were made in the OV languages. No significant effect of the PP order was observed ($F(1,33) = 1.499$; $p = .230$). The interaction between VP and PP order was also non-significant ($F(1,33) = 2.293$; $p = .139$), indicating that there is no significant difference between the natural and the unnatural languages. Although the interaction was not significant, Figure 4.9 does seem to show a pattern that is similar to the other test scores (albeit reversed as it is concerned with the number of errors).
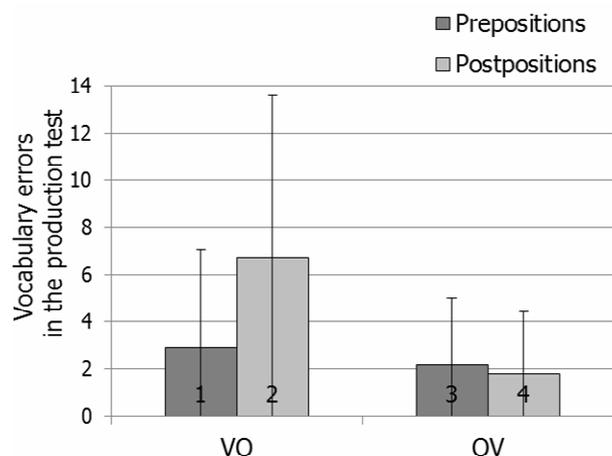

Figure 4.9. Means and standard deviations of the vocabulary errors made in the production test

5. Discussion

5.1. Vocabulary and dominant word order

As was shown in the results, the scores on vocabulary measures and the dominant word order in both simple and complex sentences were close to ceiling. For the most part, the languages were thus well-learned.

5.2. Adposition order

The adposition order was clearly the aspect of the language that was harder for participants to learn. This showed from the errors that were made in the complex sentences in the production task and the grammaticality judgment task. Note however that also for these measures, participants had fairly high average correct scores (85% for grammaticality judgments, 84% for the production task). This shows that the training that was given was adequate, allowing participants to understand and use the structure of the language. The errors that were made are therefore not simply due to the participants failing to grasp the language altogether.

For the PP order in the production task and the complex grammaticality judgments, the interaction between the dominant word order and the adposition order of the target language was significant and marginally significant, respectively. This indicates that the effects of the two variables are dependent on each other. The patterns of performance per language seemed similar for both measurements, with languages 1, 3 and 4 showing similar performance and especially language 2 showing more errors. However, the difference between language 2 and the other languages was only found to be significant for the PP order in the production test.

The (marginally) significant interactions between the two orders and the (trend towards) worse performance on language 2 can be interpreted as a combination of influence from a VP/PP bias (which is essentially a manifestation of a head-order bias, but for clarity I will use this term) and the native language. I will make two claims about the meaning of the results: firstly, the natural languages show equally good performance, which is evidence for an effect of the VP/PP bias. Secondly, scores on the unnatural languages are not equally weak, which is evidence for a combined effect of the VP/PP bias and native language influence. The first claim is most obvious, as the results are exactly in line with the predicted influence of a VP/PP bias. I will get back to this below. The second claim involves a more complex explanation, which I will provide first.

5.2.1. Combined influence on the unnatural languages and the N/PP order

As was mentioned in Chapter 2, the VP/PP universal is said to be a bidirectional implicational universal (Dryer, 2011d). This would mean that both orders imply each other to a similar extent. In that case, it would be expected that because languages 2 and 3 both differ from the native language structures on one language feature, native language influence should not cause differences between these two languages. However, I will argue here that performance on language 2 was so much worse than language 3 because language 2 was similar to English with regards to the VP order, while language 3 differed from English on that feature. In order to make this argument, I will first turn to the results from the PP-positioning to explain why there is a difference between the two orders.

The third order that occurred in the training data which is usually also consistent with the VP and PP order is the order of the PP and the noun it modifies. To avoid influence on the order of the VP and the PP, the distribution of this order was balanced, as was discussed in Chapter 3. In the production task however, this balanced input distribution was not matched in the output for all languages. As was discussed in Chapter 4, for VO languages, the ordering did not differ significantly from the balanced distribution in the input, but for OV languages, the majority of produced PPs was placed to the left of the noun. The latter is in line with the positioning that would be predicted by an OV language (Christiansen, 2000).

The prominence of left-positioned PPs in the OV languages indicates that the VP order has a steering effect on the order of the PP and the noun it modifies. This directionality has been suggested for the VP/PP universal (Dryer, 2011d), but to my knowledge it has not been proven in an experimental setting or otherwise. Note also that this is against the native ordering: 'the pig behind the box' vs. '*behind the box the pig'. The VP order thus seems to have a strong effect on the N/PP order here, which supersedes native language influence. This VP influence is further corroborated by the fact that the VP order but not the PP order had a significant effect on the results.

On the other hand, learners of the VO languages adhered to the trained distribution. Experimental studies have shown that adult language learners are able to track probabilities in inconsistent input and match these probabilities in their output (Hudson Kam & Newport, 2005; 2009). Furthermore, it has also been shown that under certain circumstances, adult learners will regularize inconsistent input (Hudson Kam & Newport, 2009; Hudson Kam & Chang, 2009). In the present experiment, learners of VO languages matched the input probability of the N/PP order, while learners of the OV languages regularized towards the order that is associated with the VP order. It is likely that these two different effects are

caused by native language influence. Native language influence in this case could be hypothesized to have two effects: it could have pushed people to use more right-positioned PPs, following their native language structure. Alternatively, the fact that the main word order, that of the VP, is similar to the native language could also make it easier for learners to focus on the exact distribution that is presented and match the probability more accurately in their output.

It seems that the latter is the correct interpretation here. This is corroborated by the fact that if we look at the individual languages, there seems to be a trend for the languages that are closer to English to match the input distribution more closely. Only in language 4 the distribution of left-positioned PPs is significantly different from 50%. This also means that if the PP order is native-like, this also has some influence on the ability to copy the input distribution (otherwise the fact that language 3 is not significantly different from 50% cannot be explained). All in all, the presence of native-like features seems to allow learners to allocate more attention to the balanced distribution of the N/PP order. Indeed, it has been argued by Hudson Kam and Chang (2009) that memory load influences the ability of learners to match inconsistent input probabilities in their output. They found that learners who learned a language in which use of determiners was inconsistent performed differently dependent on the type of production test they received. Learners who performed a normal test without reduced demands on memory load generalized the input pattern (overusing the main determiners). In contrast, learners who performed tests which reduced the demand of lexical retrieval matched the input probability closely. Presence of native-like orders in some languages in the current experiment is likely to have lowered the memory load during the production test in some way.

Altogether then, the effects of a second bias were detected in the present study. The shape of the effect follows the head-order principles that are seen in the languages of the world, in combination with a native language influence. The results are thus another manifestation of the bias towards consistent head-ordering.

These outcomes bring us back to the influence of the VP/PP bias and the native language on the acquisition of the PP order itself. As the VP order seems to be the driving order, it is likely that this order being similar to English had a stronger effect on learners than the PP order could. In this case, this led participants to be less able to acquire the PP order that was different from their native language. The driving VP order corresponding to their native language probably pushed them more strongly towards using the native PP order in language 2, causing them to make more errors than the learners in language 3. In language 3, the order

that was similar to English was not the governing order, probably making it easier to ignore the native VP/PP combination.

This may seem opposite to the findings on the N/PP order, where participants were in fact helped to match the input probability if native language features were present. However, as was discussed earlier, other studies have found that adult learners are usually well able to accurately match inconsistent input in their output, especially if no additional difficulties are present. As was argued before, the presence of native-like features will have helped learners to focus on the distribution by reducing memory load. Conversely, in the case of language 2, a combined force was likely pushing learners toward a different PP order. Firstly, the universal preference for VP/PP consistency and secondly, the driving VP order was native-like and therefore probably steering even more strongly for the native-like PP order, which coincided with the 'natural' order.

Finally, an additional native language effect was found in the significant effect of PP order on the PP order scores in the production task. Scores in prepositional languages were significantly better than scores in postpositional languages. This indicates that in production, people were more inclined to use prepositions than postpositions. This is likely to be a side-effect of the 'bad' scores on language 2 and the good scores on language 3, both resulting in more prepositions.

5.2.2. Confirmation of the VP/PP bias

The discussion so far has focused on the shape of the native language influence and its interaction with the VP/PP bias. I will now shortly discuss the evidence for the VP/PP bias itself. The presence of this bias is made evident by the high scores on language 4. Language 4 is the only language that bears no similarity to English in the main linguistic features. However, learners of language 4 obtained similar scores to those acquiring language 1. The only explanation that can hold for this is that there is indeed a VP/PP bias at work in learners' way of learning or processing the language, and that the native language influence is not the only effect that is at work here. The fact that language 4 adheres to the VP/PP bias thus makes it easy for learners to acquire it.

One might argue that the similar scores on language 3 and 4 can be interpreted in a different way. Namely, if the VP order is different from the native language, learners are not 'confused' and therefore able to learn any given order. Language 2 would then be more difficult because it had the native VP order with an unexpected PP order. However, there are some problems with this explanation. Firstly, this explanation still depends on the VP order

being the driving order which influences the PP order somehow (otherwise, again, having a non-native VP or PP should be equally difficult). More importantly, this explanation clashes with the findings on the N/PP order, which was clearly influenced by the VP order. Indeed, the VP order was able to lead learners of OV languages to create a majority N/PP order that followed the VP influence, which was not present in the input. There is no reason why the connection between the VP and the N/PP order might function in a different way from the connection between the VP and the PP order. It is therefore to be expected that the VP order will have a similar influence on the PP order. Especially because the relation between the VP and the PP and the VP and the N/PP that were found here replicate the relations these orders have in the real world data.

A result that does not seem to be easily explained is the significant effect of VP order on the vocabulary error rate in the production task. Participants in OV languages made significantly fewer mistakes than participants in VO languages. Since all participants learned the exact same word forms for each vocabulary item, there seems to be no reason for a difference in vocabulary errors. If anything, one would expect that languages which are hard to learn in terms of structure would have more vocabulary errors because learners have to allocate more attention to the structure. If this is the case, there is clearly no native language bias at work here. An adequate explanation for this does not seem to be available at this time.

To summarize, in the results of this study we see an ordering bias at work twice. Firstly, PP orders that are in line with the ordering of the VP are easier to learn. Secondly, learners will extrapolate the order of the VP to the order of the PP and the noun, despite initial balanced ordering, if native language influence is absent. Furthermore, it was found that the VP functioned as the driving base structure in these ordering effects. Native language influence was found to make it especially hard to learn a language with a similar VP order but a different PP order. However, it also makes it easier to match the input probability of the ordering of the PP and the noun.

The fact that an ordering bias is obtained despite clear native language effects shows that it is deeply engrained in the way we handle language, even showing when the language in question is a second language (or even third or fourth). The bias would thus also be resistant to extensive amounts of language contact that have an important role to play in the shaping of languages. Furthermore, the bias can already be demonstrated when people are learning a very small and simplified language. The results obtained here therefore imply that the ordering bias is very likely to be strong enough to cause the large differences between language types that are found in the real world.

To come back to the predictions that were made in Chapter 2, the interaction between the ordering bias and the native language influence that was found seems to show that prediction (b) is the closest to the data:

(b). If the native language influences performance alongside the VP/PP universal bias then the ranking of scores will be as follows:

1 > 4 > 2,3

However, it is clear that the exact shape the interaction has taken was not as expected. 1 and 4 were actually on a very similar level, indicating that the ordering bias was even stronger than was initially expected. The (unpredicted) difference between languages 2 and 3 was caused by the unexpected difference between the VP and the PP order in terms of their influence on each other.

5.3. An alternative explanation

An alternative explanation of the data and especially the good performance on language 3 may be derived from Dryer's observation (2011d) on OVS languages with prepositions. As was mentioned in Chapter 3, he remarked that the number of OVS languages with prepositions was relatively high considering the total number of OV languages with prepositions (3 and 14, respectively). Because the number of languages of this type that are currently described is so small, there is no certainty at this time whether this is due to coincidence. However, if there indeed is some universal tendency for OVS languages to ignore the VP/PP universal more easily than other language types, then the present study would suggest that this would be due to an individual bias in learners. This would be a highly interesting result as such a bias would be much harder to explain than the bias causing the VP/PP universal. The two explanations are also not incompatible: it could be that both this special bias and native language influence were favouring language 3. However, there are a number of reasons why this explanation seems less likely than the one provided above.

Firstly, as was discussed in Chapter 2, the relation between VP and PP order is an exceptionally strong and well-confirmed one, with an overwhelming majority of languages adhering to two out of the four possible combinations (94,3%; Dryer, 2011d). Secondly, as was discussed in Chapter 3, the OVS languages do adhere to the general pattern. Thirdly, out of the two unnatural types, it is in fact VO + postpositions that is more common (42 languages, against 14 for OV + prepositions). This difference could just as well be due to a

slight preference for VO + postpositions over OV + prepositions. That would make it just as likely to lead to a difference in the results of this study, which has clearly not been the case.

Altogether, while it is a possibility, and indeed an interesting one, that there is an additional bias towards OVS languages ignoring the VP/PP universal, at this point in time there does not seem to be enough empirical evidence to support such an explanation. Perhaps in the future, when more languages have been described, one of the two explanations can be chosen with more certainty. Further experimental studies could also help clarify this matter. For example, the shape of the native language effect could be verified by having speakers of postpositional OV languages perform the present experiment. If language 3 would again show better learning, there may indeed be something special about prepositional OVS languages. Alternatively, by using different word orders (such as VSO and OSV for example), it could be investigated with English speakers whether the equivalent of language 3 in this study is favoured in the same way or not. If it is, the effect is likely to be due to the combined influence of the English native language and the ordering bias. If it is not, there may indeed be more to the OVS languages and their prepositions.

5.4. Relation to other studies

The present study provides interesting additional insights to some of the studies discussed earlier in Chapter 2.

As in Culbertson and Smolensky (forthcoming), both a bias on the level of the individual learner and influence from the first language were found. The native language influence in the present study seems to have had a stronger effect however. In Culbertson and Smolensky, the native language influence was only observed in the language that was least favoured, but the other three languages were not affected. In the present study, the effect probably seems stronger because there is only one preference at work: a preference to have the heads in the VP and the PP on the same side of their clauses. The two unnatural languages are therefore both equally unnatural, allowing the native language to leave a stronger mark on the results. As was discussed above, finding this native language influence may well render the results even more interesting as it shows that the bias is strong enough to overcome the influence of a first language.

Alternatively, it might be possible that in Culbertson and Smolensky's experiment, as in the current experiment, there was a difference between the two non-harmonic orders. Recall that one of the non-harmonic languages, namely Numeral-Noun + Noun-Adjective, showed less majority order boosting than the other orders. It could be that this was also due to

a difference between the two word orders in terms of their influence on each other: one might be driving the other. Taking into consideration the results from the present study, this would have to be the Numeral-N order, which was native-like. However, the fact that exactly that type of language is also less common in the languages of the world seems to make this a less likely candidate explanation. However, it is a possibility that might deserve attention in future research.

The results discussed here were clearly not in line with those found by Cook (1988). The only time there was a significant difference between scores on the adposition orders, it was associated with an advantage for prepositions, not postpositions. The present study therefore suggests that indeed, Cook's results were an artefact of his experimental design. The present study has clearly confirmed the presence of a VP/PP bias.

The results obtained here have confirmed the findings by Christiansen (2000), in showing good performance on languages with consistent head-ordering. The bias he found thus holds in an experimental setting with linguistic, meaningful stimuli. Also, this study has shown that the bias has its effect on isolated word orders as well, not simply in overall consistent or inconsistent languages. Furthermore, the present study has shown that a native language influence can occur alongside, but will not overrule this bias.

A result that no other study has hinted on so far is the finding that the VP has a leading position when it comes to the orientation of the head order in other clauses. It shows that the bias towards consistent head ordering takes one specific order as a starting point to derive the other orders from. It should be noted that in the present study, participants received more exposure to the VP order than the PP order, which could be another explanation for this effect. However, recall here that Dryer (2011d) observed that in natural languages it is almost always the VP order that will change first and entail change in the PP order, not the other way around. The explanation that is favoured here is therefore also backed up by common processes in the languages of the world.

5.5. Relation to theories on language universals

The present study provides support for theories assuming that some type of bias on the individual level is responsible for the large-scale patterns that we see across thousands of languages today (i.e. Christiansen & Chater, 2008; Kirby, 1999; Chomsky, 1986). It is therefore inconsistent with purely historical explanations that ascribe the universal to historical processes of language change such as grammaticalization (i.e. Bybee, 1988). While these historical processes may have an added effect on the patterns we see in languages today,

the present study has shown that they cannot be the only explanation, at least not for the head-ordering patterns that were investigated here.

The present results also go against the claims made by Dunn et al. (2011) that language universals are in fact lineage-specific and caused by processes of language change. If language universals were indeed caused by specific patterns of change within lineages, the differences between the learnability of the four languages should not have occurred. This study is therefore another indication that there are still problems with Dunn et al.'s application of phylogenetic analyses to human language.

What the present study, like other artificial language learning studies, cannot determine is the exact type of bias that is involved: linguistic or non-linguistic. Following Christiansen (2000), I would like to argue however that it is most likely a domain-general bias, not a linguistic bias. As was discussed earlier in Chapter 2, there are various arguments to convince us that it is unlikely that specific linguistic biases could have evolved.

5.6. Limitations of the present study

The experiment that was done in this study was subject to a few limitations. Firstly, some compromises have been made in the design of the languages. Uncommon word orders were used to avoid strong native language effects, but it is as yet not completely certain that these orders adhere to the universal pattern (even though it is highly likely given the current typological data). Also, the incorporation of a balanced word order made the universal languages more consistent in their head-ordering than the inconsistent languages. Finally, and of a more practical nature, there was a fairly limited number of participants due to constraints on time and resources. However, there are no unexpected or otherwise inexplicable results that lead back to any of these limitations. If anything, the decision to include a balanced ordering has led to an additional highly interesting and unprecedented result which has proven to be useful for the interpretation of the target data. Furthermore, the small number of participants which still led to significant results is an attestation to the strength of the bias that was found. However, as was mentioned in Chpater 4, there were a few trends to be observed which confirmed the more general patterns found in the data, but which did not reach significance. It is possible that that would have been the case if a larger group of people would have participated, which would have strengthened the results.

5.7. Suggestions for future research

A few suggestions for future research were already provided above, with regards to testing a possible special bias for prepositional OVS languages. Especially recruiting participants with different native languages for an experiment such as the one reported on here would be an interesting avenue of research. It could help to confirm or reject the interpretation of the results that were found in the present study. Another possibility might be to investigate the effect of the biases found here in chains of transmission, such as those used in iterated learning experiments. It has been suggested that underlying biases are weak, but that their effects become larger over time when language is transmitted, as small changes to the languages accumulate (Smith & Wonnacott, 2010; Reali & Griffiths, 2009). It would be interesting to see if in such an experiment the VP/PP bias would be able to push learners of an inconsistent language to gradually make the language adhere to the VP/PP bias. Whether the native language effect is sustained in such a paradigm is also an interesting question. Finally, to confirm the driving force of the VP order, it would be interesting to conduct an experiment which involves an equal number of exposures to the VP and the PP order, for example by including simple statements involving only an NP and a PP.

Investigating biases for other word orders could be another interesting step. Because the universal patterns across languages are not equally clear and strong for each order, it will be interesting to see if the bias towards certain combinations is also stronger or weaker than for others. Another interesting avenue for future research may be to further investigate the importance of the VP order in determining other orders, and possibly, what effect arises if no VP is present or if the VP order is balanced. This could tell us more about the prominence of this order and the way it is used by learners.

Finally, the present study has shown that with a relatively small amount of training, interesting results can be obtained. Tily and Jaeger (to appear) stress that time is a problem with artificial language learning studies as they usually take several sessions. However, like Culbertson and Smolensky's study, the experiment used for this dissertation took place in one session. Furthermore, participants took between 35 and 45 minutes to complete it. Depending on the linguistic features under investigation, the language can thus be sufficiently simple to be learned in a short period of time. Also, the present study has shown that an important measure to save time, namely limiting vocabulary, need not compromise the acquisition of the grammar.

6. Conclusions

The present study has shown that the universal pattern that is found in the combination of VP and PP order is caused by a bias on the individual level of the learner. This was shown by the good results on both natural languages in the experiment, whether the target structures were similar to English or not. This bias for consistent head-ordering was also observed in the effect of the VP order on the order of the noun and its modifying PP. This became apparent through the preference for left-positioned PPs for the N/PP order in OV languages. In addition, it was shown that it is in fact the VP that exerts this influence on the other word orders, as the effect of the VP was significant, but not the effect of the PP.

Alongside the effects of these biases, native language influence was observed. This native language influence became apparent in the more accurate matching of the balanced input distribution of the N/PP order in VO languages as opposed to OV languages, and the more accurate matching of this distribution in all three languages with native language features. Furthermore, there was a difference in the two unnatural languages, with the language that shared the VP order with English showing worse performance. It was argued that this was due to the fact that the VP has the leading influence on the other orders.

The present study thus lends support to theories which suppose that innate biases are at the basis of the universal patterns we see in natural languages.

Word count: 14,723

References

Bybee, J. L. (1988). The diachronic dimension in explanation. In J. Hawkins (ed.), *Explaining Language Universals* (pp. 350-379). Oxford: Blackwell.

Christiansen, M.H. (2000). Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In J.L. Dessalles & L. Ghadakpour (eds.), *Proceedings of the 3rd International Conference on the Evolution of Language* (pp. 45-48). Paris: École Nationale Supérieure des Télécommunications.

Christiansen, M. H. (1994) *Infinite Languages, Finite Minds: Connectionism, Learning and Linguistic Structure*. PhD thesis, University of Edinburgh.

Christiansen, M.H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5), pp. 489-558.

Christiansen, M. H. & Devlin, J. T. (1997) Recursive inconsistencies are hard to learn: A connectionist perspective on universal word order correlations. In: M. G. Shafto & P. Langley (eds.) Proceedings of the 19th Annual Cognitive Science Society Conference (pp. 113-18). Mahwah, NJ: Lawrence Erlbaum.

Christiansen, M.H. & Ellefson, M.R. (2002). Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In A. Wray (ed.), *The transition to language* (pp. 335-358). Oxford: Oxford University Press.

Chomsky, N. (1986). Knowledge of language: Its nature, origins and use. Westport: Praeger Publishers.

Chomsky, N. (1980). On cognitive structures and their development: A reply to Piaget. In M. Piattelli-Palmarini (ed.) *Language and learning: The debate between Jean Piaget and Noam Chomsky* (pp. 35-52). Cambridge, Massachusetts: Harvard University Press.

Comrie, B. (1981). *Language Universals and Linguistic Typology.* Oxford: Basil Blackwell.

Cook, V.J. (1988). Language learners' extrapolation of word order in micro-artifical languages. *Language Learning*, 38(4), pp. 497-529.

Croft,W., Bhattacharya, T., Kleinschmidt, D., Smith, D. E., and Jaeger, T. F. (to appear). Greenbergian universals, diachrony and statistical analyses. *Linguistic Typology*.

Culbertson, J. & Smolensky, P. (forthcoming). Testing Greenberg's Universal 18 using the Mixture Shift Paradigm for artificial language learning. To appear in *Proceeding of NELS 40*.

Dryer, M.S. (2011a). Order of object and verb. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online* (chapter 83). Munich: Max Planck Digital Library. Retrieved 7 June 2011. Web site: http://wals.info/feature/83A.

Dryer, M.S. (2011b). Order of adposition and noun phrase. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online* (chapter 85). Munich: Max Planck Digital Library. Retrieved 7 June 2011. Web site: http://wals.info/feature/85A.

Dryer, M.S. (2011c). Order of subject, object and verb. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online* (chapter 81). Munich: Max Planck Digital Library. Retrieved 7 June 2011. Web site: http://wals.info/feature/81A.

Dryer, M.S. (2011d). Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase. In M.S. Dryer & M. Haspelmath (eds.) The World Atlas of Language Structures Online (chapter 95). Munich: Max Planck Digital Library. Retrieved 29 July 2011. Web site: http://wals.info/chapter/95.

Dryer, M.S. (2011e). Order of subject and verb. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online* (chapter 82A). Munich: Max Planck Digital Library. Retrieved 2 August 2011. Web site: http://wals.info/chapter/82A.

Dryer, M.S. (1992). The Greenbergian word order correlations. *Language*, 68(1), pp. 81-138.

Dunn, M., Greenhill, S.J., Levinson, S.C. & Gray, R.D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, pp. 79-82.

Evans, N. & Levinson, S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, pp. 429-492.

Greenberg, J.H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (ed.), *Universals of Language* (pp. 58-90). Cambridge, Massachusetts: The M.I.T. Press.

Hawkins, J.A. (1983). *Word order universals*. New York: Academic Press.

Hudson Kam, C. L. & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(3), pp. 815-821.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59, pp. 30-66.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1, pp. 151-195.

Kirby, S. (1999). *Function, selection and innateness: The emergence of language universals*. Oxford: Oxford University Press.

Kirby, S. (1998). Fitness and the selective adaptation of language. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (eds.) *Approaches to the Evolution of Language: Social and Cognitive Bases* (pp. 359-383). Cambridge: Cambridge University Press.

Lupyan, G. & Christiansen, M. H. (2002). Case, Word Order and Language Learnability: Insights from Connectionist Modeling. In W.D. Gray & C.D. Schunn (eds.) *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 596-560). Mahwah, NJ: Lawrence Erlbaum.

Reali, F. & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111, pp. 317-328.

Smith, K. & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116, pp. 444-449.

St. Clair, M. C., Monaghan, P. & Ramscar, M. (2009) Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science*, 33, pp. 1317-1329.

Tallerman, M. (2005). *Understanding Syntax, Second Edition*. London: Hodder Arnold.

Tily, H. & Jaeger, T. F. (to appear). Complementing quantitative typology with behavioural approaches: Evidence for typological universals. *Linguistic Typology*.

Tily, H. J., Frank, M. C. & Jaeger, T. F. (2011). The learnability of constructed languages reflects typological patterns. *Proceedings of CogSci 2011* (pp. 1364-1369).

WALS online (2011a). Combined features: Order of object and verb and Order of adposition and noun phrase. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Retrieved 7 June 2011. http://wals.info/feature/combined/83A/85A

WALS online (2011b). Combined features: Order of subject, object and verb and Order of adposition and noun phrase. In M.S. Dryer & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Retrieved 07 June 2011. http://wals.info/feature/combined/81A/85A

Wonnacott, E., Newport, E. & Tanenhaus, M.K. (2007). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, 56, pp. 165-209.