The acquisition of "optional" movement
Zuckerman, Shalom

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2001

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

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

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 2

Language Acquisition

1. Theory of parameter setting

The Principles and Parameters framework offers a general picture of how language is acquired. The child, equipped with an innate form of grammar, faces the task of setting the parameters of U.G. to the correct values of the target language. The two sources of information the child uses in this task are the fixed universal principles, which direct the child to consider only grammars which belong within the range of possible human languages, and the environmental input, which gives her positive evidence with respect to the setting of parameters in the target language.

If we said earlier that the task of linguistic theory is to solve Plato’s problem, to explain how children acquire language so rapidly, then the task of language-acquisition theory is the opposite, and contrasts with Plato’s problem, which is the “logical problem” of language acquisition. Hyams (1988) calls this the “developmental problem” of language acquisition: once a model is presented that successfully addresses Plato’s problem and explains the relative ease with which language is acquired, one must account for the fact that acquiring a language does not happen instantly. If the only thing the child needs to do is set the value of the parameters, the data available to her through the input should, in principle, be sufficient. Yet observations of child speech, as well as children’s comprehension abilities, show that children go through stages in which their language is not adult-like. During these stages, children produce sentences that are ungrammatical in the target language, and, from an adult perspective, they interpret sentences incorrectly.

Some explanations of the developmental problem are obvious. For example, elements of the lexicon such as words and morphemes must be assumed to be learned item by item. This, of course, takes time. Furthermore, some grammatical rules are dependent on specific properties of these lexical elements; therefore such rules, although present in the child’s grammar, cannot be implemented before the relevant lexical items and their specific properties are acquired.

Other explanations are less obvious. For example Borer and Wexler (1987) raise the possibility that there is a maturation of some of the grammatical principles. White
(1981) and lately Wijnen (1998) propose a distinction between ‘input’ and ‘intake’ and discuss the possibility that the child does not make actual use of all the sentences she hears, from one reason or another. Such explanations, although possible, are not necessary and depend on heavy empirical evidence to support them.

Clark and Roberts (1993) pour some cold water on the conception that the P&P model offers a full explanation to the relative ease with which language is acquired. They define the term hypothesis space to include all possible grammars the child might consider, and calculate that if we assume U.G. contains 30 binary parameters, which is a conservative assumption, then there are $2^{30} = 1,073,741,824$ possible grammars for a child to consider. Even if we assume the child tests one possible grammar per second this will lead, in the worst case, to 34 years before the correct combination of parameter settings is acquired.

Putting the points above together, the picture that emerges is that the theory of parameter setting (that is, the theory of language acquisition under a P&P approach) must formulate principles that will account for the rapid and efficient manner in which parameters are set, but at the same time explain the intermediate, non-adult-like, stages through which children go.

In this chapter I wish to propose such a principle, which on the one hand facilitates successful acquisition and on the other hand is responsible for non-adult-like behavior in the intermediate stages. This principle is a familiar one: the principle of economy. I will claim that the principle of economy which is designed as a tool for selection among equally optimal derivations, is adopted by the learner as a tool for selection among equally optimal grammars.

2 Assumptions and guidelines

Recall that the minimalist program, under the heuristic of external minimalism, calls for a minimization of assumptions that are made to account for empirical phenomena. This heuristic also applies to the field of language acquisition, which attempts to explain the empirical observation that children acquire their mother tongue in an efficient, successful manner while going through non-adult-like intermediate stages. In the following subsections I will consider several relevant assumptions made in the literature, and try to show how they are justified under such a requirement for minimalism.
2.1 No negative evidence

It is commonly accepted that negative evidence is not available to children (e.g., Wexler (1978), Baker (1979), Wexler and Culicover (1980), Pinker (1989). Although children are sometimes exposed to corrections of their errors, these seem to have no impact on their progress. Furthermore, it has been claimed that such negative evidence is inconsistent, since some parents correct their children while others do not, and that in most cases children interpret corrections as relating to content and not to structure. These empirical observations lead to the assumption that a theory of language acquisition must explain how language is acquired without the availability of negative evidence. Note that this is not an assumption of a principle of language acquisition but rather an empirical description of the way language is acquired in practice. I will adopt the assumption of no negative evidence and assume that language must be shown to be acquired on the basis of positive evidence alone.

2.2 No memory for past input

A second assumption I will adopt is the assumption that the child does not remember structures that appeared in the input she receives and thus that such past input strings are not available when an hypothesis for an intermediate grammar is made. This means that at any given input string, the child values the current hypothesis parameter settings she holds against this specific string. It is possible, of course, to compare this new string with other strings generated by the child’s current grammar (cf. Braine 1971). Although the question of whether the child has memory for past input is empirical, it is hard to imagine what type of evidence can be brought to support or refute this claim. Therefore, the inclusion of such an assumption should be based on considerations of simplicity. As long as a model that assumes no such feature is able to account for the empirical facts of language acquisition, it should be preferred. I shall therefore assume that past input strings are not available for the learner and that grammar evaluation is done at each given instance based on a single input string.
2.3 The subset principle

The subset principle (Berwick 1985, Wexler and Manzini 1987) can be seen as a corollary of the two assumptions made above. This principle states that language acquisition progresses in a fashion such that any intermediate grammar assumed by the child is always a subset of the target grammar. The idea is that without negative evidence once a superset grammar is assumed, no positive evidence is able to restrict this grammar to the subset target grammar, since all structures generated by the target grammar can also be generated by the intermediate grammar.

Under the assumption that an intermediate hypothesis is revised only when an input string cannot be parsed, a superset intermediate hypothesis will never be revised and the target grammar will never be acquired. Such reasoning has led researchers such as those above to assume the subset principle as an innate principle of language. Note that the assumption of no memory of past input is necessary for the subset principle. If the learner remembers all past input strings, then one can imagine a situation in which an intermediate superset hypothesis is replaced with a subset hypothesis when the learner realizes that another simpler hypothesis can account for all past input strings. Nevertheless, once the assumptions of no-negative-evidence and no-memory-of-past-input are made, the subset principle, in my view, follows automatically. However it is not necessary to adopt Berwick’s (1985) or Wexler and Manzini’s (1987) proposals for an innate principle designed specifically to disallow a superset hypothesis. It is sufficient to assume that principles of language acquisition must be formulated in such a way that superset hypotheses do not arise. It will be proposed in this chapter that the principle of economy obeys this requirement since it blocks the possibility of one type of a superset hypothesis namely, the hypothesis of optional movement.

Notice that under the assumptions of the minimalist program, there are no subset-superset relations between a rule of obligatory movement and a rule of obligatory no-movement or covert movement. A child who mistakenly assumes movement where it is forbidden will receive positive evidence for her error and a child who assumes non or covert movement where it is necessary will also receive such evidence. (Recall that under the MP, movement is always either motivated or blocked by economy. This means, for example, that subset-superset relations

---

11 Note that this can be done without the presence of a new input string, under the assumption that the learner constantly searches for a simpler hypothesis that will explain the past input strings.
between a grammar that has subject-auxiliary inversion and a grammar that has no such inversion (cf. Weinberg 1990) cannot be formulated; in the former the movement is obligatory and in the latter it is forbidden. On the other hand, as will be claimed later, a grammar that allows for optional movement is indeed a superset grammar with respect to a grammar that does not allow optionality. We conclude therefore, that the subset principle has an important role in language acquisition theory, although not necessarily as an explicit principle. Rather it is a heuristic to guide us when we make assumptions about principles of language acquisition.

2.4 Poverty of the stimulus

Recall that we said that one of the main goals of linguistic theory as well as of the theory of language acquisition is to offer a satisfactory answer to Plato’s problem. The problem of the poverty of the stimulus (cf. Crain 1991) is, in one respect, identical to Plato’s problem: how are children able to produce an infinite number of sentences based on a finite number of input sentences? But the input children receive is also poor in another sense. Apart from being quantitatively poor finite and moreover simplified and repetitive, evidence has shown that it is also qualitatively poor. The input to children has been shown to include many cases of incomplete utterances, pauses and a ‘change of mind’ in a middle of an utterance, as well as simple performance errors. To those one should add ungrammatical input from other children or from non-native adult speakers. As a result the child is said to be exposed to a significant number of utterances that might impair the acquisition process. In order to overcome this problem, the child should be endowed with the ability to filter-out such input strings. Moreover, many languages manifest differences resulting from different registers, ‘styles’, dialects, etc. This can be seen as another form of ‘input corruption’ for which the child must, instead of filter-out irrelevant strings, discriminate between two different ‘contexts’ register, style, dialect.

The conclusion is therefore, that the theory of language acquisition must propose principles that function as a filtering mechanism for irrelevant input strings.
2.5 The LAD.

The concept of language acquisition device (LAD) was introduced to linguistic theory (Chomsky 1965) as symbolizing the procedure through which language is acquired. Under the common view, the LAD is said to be the connecting link between the primary data, U.G. and the final state – the target language. Following our guideline of questioning our hypotheses, we must question the necessity of the LAD itself. By that I do not mean to question the innate basis of language that I assumed in the previous chapter, but rather to ask if there are principles by which the LAD operates that are different from the principles by which the computational system operates. Put another way, the question is whether children are endowed with rules specifically designed for acquiring rules or setting parameters. Such principles were proposed in the literature (e.g. in Wexler and Culicover, 1980). If it could be shown that the principles of acquisition are identical to, or at least share a common basis with, principles proposed for the computational system independently of acquisition considerations, then the notion of the LAD, as distinct from U.G. would become redundant. It is therefore clearly a minimalist goal to show that principles of acquisition depend on principles of computation. It is not my purpose to claim that this is the case and that there are no principles unique to acquisition, but rather to adopt a guideline that prefers existing computational principles over unique acquisition principles.

3 Optionality in child language

In the previous chapter we discussed the phenomenon of apparent optionality in the adult grammar. We adopted a direction that objects to true optionality, assumes that U.G. reflects correspondence uniqueness between meaning and sound, and leads to a deterministic relation between the numeration and the PF structure. The impossibility of true optionality is assumed to be inherent in U.G and thus common to all languages. A corollary of this assumption is that children must reject optionality as well, as universal principles need not be learned. We further claimed that the assumption of optionality in the target grammar is devastating for the learner as it consists of a superset assumption from which one cannot revert. This point was observed in Wexler (1978), who points out that no positive evidence can cause a child to revise an overly general optional rule, and by Clark (1987), who
proposes an inherent principle of contrast that is available to children from the earliest stages. The conclusion is thus that child language should not reflect cases of optionality.

On the other hand, several phenomena reveal what seems to be optionality in child language. Moreover, children seem to manifest optionality in cases that require obligatoriness, thus they seem to be turning an obligatory rule into an optional one. Examples of such phenomena are the optional omission of subjects by speakers of non-pro-drop languages that require an overt subject known as the null subject phenomenon (Hyams 1986), the optional production of non-finite verbs as main verbs (the optional infinitives phenomena, Wexler 1994), and so forth. The fact that children produce what look like optional structures, seems to contradict the assumption we made above -- specifically the subset principle-- in the same way that the apparent optional structures presented in the previous chapter seemed to contradict the assumption of no optionality in adult language. This apparent contradiction led researchers such as Van Kampen (1997) to abandon the subset principle. I wish to take the opposite position and object to the claim that optionality in child language is real, in order to maintain the principles that do not allow it, such as the subset principle and the principle of uniqueness, as well as the principles of economy. An empirical support for the claim that child language does not reflect true optionality comes from De Haan (1987) (and lately confirmed in Wijnen 1996), who observes that the types of verbs used in root-infinitive structures in Dutch are different from the ones used in finite structures. This shows that children do not hold a rule with an optional value but rather apply different rules to different verb classes. As the phenomenon of null subjects was shown to be tightly related to the use of infinitives (Wexler 1999), one can speculate that in this 'optional' behavior the children apply in fact different obligatory rules to different verb classes. Following this observation, I will adopt the hypothesis that optionality in child language is not the result of children including an optional rule in their grammar but rather an application of several related obligatory rules to different structures (different numeration sets).

Another argument for the claim that children avoid optionality is the following: in a previous section we have seen that the input to children might include structures that must be either ignored (for example broken sentences, false starts, and performance errors or classified as belonging to a different style, register, dialect, etc). These cases result in contradictory input. Children must be equipped, then,
with a principle of selection to help them in the task of ignoring and classifying any contradictory structures that might impair the acquisition process. Moreover, Clahsen (1991) observes that this problem is not restricted only to cases of corrupted input, but might also result from children’s misanalysis of ‘proper’ data: the child might fail to discriminate certain properties of structures that appear in her input, and incorrectly conclude that the input is contradictory. Clahsen speculates, for example, that a German-speaking child might fail to distinguish finite from non-finite verbs and perceive contradictory input with respect to the direction of the head of the VP projection. Failure to distinguish matrix and embedded clauses, interrogatives and declaratives, etc., might also lead to contradictory input. Now, if we assume that true optionality is acceptable to the child, we risk the possibility that these cases of contradictory input will be also classified as legitimate optional structures. Such an assumption is bound to impair the child’s ability to make the proper discrimination in the future. In other words, the child must be equipped with the tendency to keep looking for evidence that will enable discrimination between contradictory evidence. A rejection of the possibility of optionality results in this very tendency, as cases of apparent optionality are rejected and the child is forced to continue searching for evidence that will either allow for different classification of the two contradictory options or lead to ignoring one of the options.

The argument in the previous paragraph leads to an observation of another problem inherent to the acquisition process, with respect to the discrimination between I-language and E-language made in the previous chapter. The task of the language-learning child is to set the parameters of his I-language. However, the input through which the child must fulfill this task does not represent a sample of an I-language but rather of the E-language. As the I-languages of different speakers who contribute to the input might be different from one another, the child must be able to incorporate these different I-languages into a single I-language. In this case, therefore, the need for a selection rises again calling for a principle that will lead the child to reject contradictory input and adopt only one of the options.

To conclude, in this section I have claimed that we should expect learners to reject the possibility of optionality. I have considered arguments from the point of view of learnability, and shown that an assumption of optionality is hazardous to the

---

12 I use the term E-language here to refer to the collection of all well-formed structures in the collection of I-languages of all speakers in a linguistic community. Although this is an abstract notion since there is an infinite number of such well-formed structures, the notion of ‘a sample of the E-language’ is a concrete one and is different from the notion of I-language.
acquisition process. I have also presented empirical evidence showing that phenomena reflecting apparent optionality in child language represent an application of different rules to different numeration choices. In the next section the principles of economy are revisited, this time from the perspective of child language. I will claim that in acquisition, as in computation, economy functions as a selection process operating on contradictory input.

4 Economy in Child language

Before the introduction of the Principles and Parameters model (Chomsky 1981) and its application to language acquisition (Hyams 1986), the procedure of acquisition was assumed to be based on hypothesis evaluation. Specifically, the child makes a hypothesis, constrained by the set of possible human languages, with respect to the set of rules that constitute the target grammar, then evaluates this hypothesis against the input data according to some evaluation metric. Notions of simplicity and economy were raised then (e.g. Chomsky 1965) as possible metrics for the evaluation of grammars. Under the P&P model, however, none of this is necessary. The child does not have to evaluate grammars, but rather has only to set, at each point, a specific parameter based on a specific piece of input. Under this approach, comparing or evaluating a hypothesized grammar is reduced to setting a value for a specific parameter against a single string from the input. As pointed out in Chomsky (1998) the notion of economy seems to be irrelevant for such a procedure (but see Clark and Roberts, 1993 for incorporating the evaluation metric into a parameter setting approach). In the current study the notion of economy is reintroduced for the process of language acquisition under a minimalist approach. Recall that we adopted the ‘global’ interpretation of economy of derivation, viewing it as operating on a set of alternatives rather than on a single candidate. The notion of a set of alternatives must be well defined for global economy to be meaningful. Under the old concept of evaluation metric, the set of alternatives was not clear (it was essentially all possible grammars) and as a result notions of economy, complexity and simplicity were vague. However, once the notion of economy was sharpened in the Minimalist Program, defined as relevant only to alternatives stemming from a single numeration, and clarified with respect to the measures by which one of the alternatives is ruled more economical, it is possible to reconsider the role of economy in acquisition.
Several authors have observed that economy plays a role in language acquisition in the sense that movement operations are more vulnerable than non-movement operations. I will review three such proposals.

Clark and Roberts (1993) propose a model of learnability and language change that is based on the model of genetic algorithms. This model is said to begin with a random setting of parameter values and continue through evolution-like procedures of selection of better candidates. This model incorporates within the notion of the evaluation metric, which evaluates the hypothesized grammars and selects the ‘fittest’ among them, a component of economy. The primary factor by which hypothesized settings are valued is their ability to parse the input. To deal with cases in which two or more hypothesized settings score equally on the parsing scale, the authors propose a compactness factor that determines which of these alternative hypotheses is to be selected. They describe compactness as follows:

The learner will, all else being equal, prefer hypotheses that lead to more compact representations. Compactness, here, can be defined in terms of such factors as the number of nodes required to cover the input string, the length of the chains associated with arguments and operators, or both.” (pp. 313)

Clark and Roberts bring empirical evidence from language change to support their assumption that hypotheses that score higher on the compactness factor are preferred. The authors’ definition of compactness can be seen as a pre-minimalistic notion of economy as it is instantiated in acquisition. Notice that this definition is a global one, since hypotheses are ranked as compared to each other and a specific hypothesis cannot be said to be compact or not compact based on its properties alone.

A second proposal with an economy flavor is made in Platzack (1996). Platzack proposes, already within a minimalist framework, that learners hold an initial hypothesis based on which they evaluate input. This is the hypothesis of ‘weak’ values for all features. This initial hypothesis, in combination with economy, contributes to the emergence of early stages in language acquisition in which children fail to perform movement, and to the lack of stages in which children assume movement (i.e., strong features) when the target grammar does not support such an assumption. Platzack reviews errors in children’s Swedish that reflect failure to perform verb-movement. This proposal can be described as a ‘local’ approach with respect to the global/ local distinction. According to Platzack’s proposal, any
A third proposal with an economy element is made in Van Kampen (1997). According to this proposal, children’s early stages are characterized by an attempt to decrease the discrepancy between the LF and the PF representations. As the main source of such discrepancy is movement, the prediction that emerges from this proposal is that children will fail to perform operations that involve movement since these represent a larger discrepancy from the LF structure, and instead will produce structures in which elements are in or closer to their base position. Van Kampen appeals directly to the notion of economy of derivation and claims that in the acquisition process, economy serves the children as a performance-strategy leading to the preference of more economical structures in the early stages. Van Kampen brings evidence from early pivot-structures, Do insertion, Wh-drop and medial-Wh phenomena in Child Dutch to support the claim that the early stages are characterized by the use of non-adult structures that are more economical than the target ones. With respect to our local/global distinction, Van Kampen’s proposal falls with Platzack (1996) on the local side, claiming a general problem with movement rather than a preference towards no-movement when both options are available.

We have seen three recent proposals that incorporate a notion of economy within their description of the acquisition process and tried to distinguish among them based on the local/global distinction. With respect to economy in computation (reviewed in the previous chapter) the local approach claims that economy principles mark a structure as economical, and therefore convergent, without comparing it to other alternatives. The global approach marks a structure as economical only with respect to the other alternatives stemming from the same numeration, which would have converged had economy not played a role. In our transfer of economy principles to the domain of acquisition, the local approach predicts a structure to be more economical, and thus easier to acquire, without relation to other structures present in the child’s input. The result is that the local approach predicts that any structure reflecting movement will be problematic for
children, while the global approach predicts that structures will be seen as economical and preferable only with respect to other alternatives (stemming from the same numeration and convergent if economy is ignored) present in children’s input. The proposal of the current dissertation, to be presented in the next chapter, takes the global approach, and therefore applies specifically to the cases of apparent optionality in the target language. The idea is that parameters are easy to set as long as the triggering data is clear and unanimous. However, in cases where the triggering data appears contradictory, the task becomes more difficult. In these cases a selection mechanism is needed in order to select one structure from the seemingly contradictory alternatives as a trigger for the parameter. According to the current proposal, it is only in these cases of contradictory input that economy principles apply.

Although the distinction between the local and the global approaches seems to be theoretically clear, it is not so clear whether one can point to empirical evidence that will distinguish between them. For example, if a specific obligatory movement operation is shown to be present in child-language from the earliest stages, this seems to go against the prediction of the local approach that any movement operation will be problematic; however it could be claimed that the relevant parameter is acquired prior to the onset of speech (i.e. very early parameter setting, cf. Wexler 1999). On the other hand, if children are shown to have problems with an obligatory movement operation that does not have a clearly more-economical counterpart in the input, this would contradict the global approach. However, it could be claimed (as will be claimed in chapter 6) that due to a misanalysis of the input, children might view a structure that is not related to this problematic structure as related to it and thus think that they do have a more economical alternative and opt for it. These questions will be discussed throughout the dissertation and will be revisited in the general discussion.

The current proposal, therefore, assumes that children employ the principles of economy, which are independently assumed for computation, in the task of acquisition. As in computation, economy is used by learners to choose from alternative candidates, of which only one must surface as a trigger for a specific parameter. Notice that the description of the principles of economy as principles that guide acquisition fulfils many of the requirements we listed above for a

---

13 The notion of LF in Van Kampen (1997) is different from the usual (cf. Huang 1992, Chomsky 1995) LF. Van Kampen assumes a procedure of reconstruction in LF, leading to an LF representation that represents
principle of language acquisition. Economy prevents a superset assumption from emerging, it serves to classify and filter input, and it accounts for possible non-adult like stages in which the child mistakenly applies economy to two structures that are not true competitors, since they do not in fact stem from the same numeration. With respect to minimalism, the current proposal answers the requirement of minimizing assumptions by assuming one principle plays a role in both computation and acquisition.

J. Uriagereka, in his book *Rhyme and Reason*, puts the following text in the mouth of the ‘Linguist’\(^{14}\):

... the sort of LAD I discussed - which is capable of picking out the simplest analysis of messy data, according to a linguistic metric - may have the problem solving capacity that U.G. implies, for the computational system. This is assuming that the LAD adopts given structural analyses because they involve a hypothesized linguistic computation of the shortest kind that is, the shortest kind that’s compatible with input data. I might add that this isn’t a necessary assumption, but it’s a natural one: provided that the LAD is programmed with all the information that U.G. carries, its also capable, in principle, of using that information to analyze the messy data that it gets. (pp. 78)

And actually, this is what I would like to say.

5 Conclusion

In this chapter, the notions of optionality and economy are examined from the perspective of child language. Reviewing the basic requirements of the field of language acquisition, it was concluded that theory of language acquisition must include principles to explain not only the speed and accuracy with which language is acquired but also the non-adult like stages through which the acquisition process goes. Furthermore, we showed that these principles should allow for the filtering of the irrelevant data in the input and should explain the choices children make when the input includes evidence contradictory to the target grammar. We speculated that a model of language that includes similar principles for computation and

\(^{14}\) The book is presented in the form of a dialogue between a linguist and an interlocutor who explore the basics of minimalist syntax together.
acquisition is superior to one that includes different principles for these two tasks. With respect to optionality and economy, the following claims were made:

- **Optionality is undesired from the perspective of learnability.** Children avoid the assumption of optionality as a characteristic of the target grammar; therefore it will not be a characteristic of the intermediate grammars. This claim is supported by the subset principle, by the claim made in the previous chapter that optionality is not allowed in the adult grammar, and by empirical evidence showing that optionality in child language is not real.

- **Economy plays a role in the acquisition process.** Child language is characterized by less movement when compared to the target grammar. Structures that include movement seem to be more problematic for children than structures that reflect the base order.

- **Economy is assumed here to be of a global nature;** children use the principles of economy to decide among related alternatives in their input in the same way that the computation system uses them to choose between related derivations for convergence. Thus, economy in child language applies only when children face structures in their input that seem to be optional.

In the next chapter we combine the conclusions of this chapter with those made in the previous chapter to create a general proposal with respect to seemingly optional movement operations and their acquisition. This proposal will lead to empirical predictions that will be tested in the second part of the dissertation.