3

Ard: Non-linear models and theories

Ard, guth is saorsa  
Ard, tha sinn gluasad  
Gu h-ard theid sinn suass  

Ard (High) by Runrig from the CD Amazing things

3.0 Introduction

High, Ard in Scots-Gaelic, is a place somewhere above. For a child to get high, i.e. to grasp grammatical knowledge, he has to move, he has to change the low into a high. Standard psychological and statistical methods, and methods for the analysis of empirical data assume a linear underlying form of change (e.g. linear regression) and a unimodal distribution of scores. Suppose that at least some parts of (language) development are basically non-linear or even discontinuous, and stages (i.e. a multimodal distribution in empirical data) exist, what models can be used to prove that the development of child language is a non-linear, or even discontinuous process? This question is answered in this chapter.

The crucial discussion in the previous chapter is the one on discontinuity and continuity. This discussion has not been settled yet, due to the absence of criteria and a meaningful definition of discontinuity. In chapter 2, I offered a working definition of discontinuity, which is:

There is a qualitative (structural) change in development in such a manner that there is no structural linkage (i.e. after the change there is a structure that was not there before) with earlier behaviour, and this change is reflected by the lack of smoothness of a growth curve.

This definition is not very useful if applied to language development. For example, what does, in empirical terms, ‘no structural linkage’ and ‘a lack of smoothness’ mean? Thus, a discussion on discontinuity is more or less futile because of a lack of a conclusive definition and method. However, a new research method has been introduced in (developmental) psychology in roughly the last twenty years. This method aims at providing criteria for the study of discontinuities, and it embraces the concept of non-
Chapter 3

linearity. Non-linear models proved to be fruitful in (developmental) psychology, and they are discussed in this chapter.

In section 3.1, continuous linear and non-linear models and theories are discussed with a special interest in the application to (developmental) psychology. These models and theories are used to solve developmental problems. I shall explain the non-linear equations of these models.

Some theories of human development, for instance Piagetian theory, assume stage wise (or discontinuous) development. However, a model and theory of such development with unambiguous criteria has been absent. Catastrophe Theory has been developed in order to demonstrate discontinuities in development (section 3.2).

In the summary of this chapter, the basic question of how to explain change, in particular non-linear change, is reformulated, with the stress on quantitative aspects (instead of qualitative, i.e. linguistic, ones).

3.1 Continuity: linear and non-linear models

Introduction

Non-linear models began to flourish when Lorenz and his collaborators ran a series of experiments. These experiments were carried out in the sixties, and they involved extensive calculations of weather forecasts. Due to some small accidental changes in initial values, Lorenz found enormous changes in the end values. He and his collaborators ran these calculations again because they thought that they had made a mistake somewhere. But they had not. Their calculations lead to the discovery that small changes in the initial state lead to big differences in the end and that weather forecasts only have a limited predictive value. This discovery, ‘small changes may lead to big differences’, is the ‘butterfly of Lorenz’: when a butterfly in Brazil strikes its wings, it may eventually cause a thunderstorm in Tokyo.\(^1\)

Lorenz’s finding has indeed caused some strong winds. However, they did not take place in Tokyo or any other particular place, but they affected a host of scientific research programs (e.g. patterns in fluctuations on the capital market can be explained by using non-linear paradigms; cf. Peters, 1991). In psychology, the application of catastrophe (from the Greek kata for ‘down’ and strophé for ‘turn’) theory by Zeeman (1976) has been propitious for the study of sudden changes in the behaviour of animals (i.e. dogs). In developmental psychology, Thelen (1989) and van Geert (1991) are strong

---

\(^1\) An introduction into chaos and non-linear processes is given by Gleick (1987) and Stewart (1989).
advocates of applying dynamic systems models to the development of language and to motor development. The application of these models was also inspired by the fact that traditional statistical methods did not suffice (i.e. the failure of the linear paradigm; Peters, 1991).

Two important issues are discussed in order to be able to apply these models to language development. First, I discuss dynamic systems models (DSM) and dynamic systems theory (DST) that have become popular in a very short time. Second, dynamic system models are built up from difference or differential equations. I discuss the elements of the equations that constitute a mathematical model of cognitive development (van Geert, 1991), and the results of existing modelling.

The reason to use dynamic systems models is that they can be used to describe behaviour in biological and chemical processes that show chaotic behaviour, that is, behaviour that does not follow smooth paths along the time continuum. By using dynamic systems models in the form of differential equations, scientists were able to explain the behaviour of biological dynamic interactions like predator-prey-models. In fact, all behaviour which has a time component and all behaviour which cannot be explained by standard linear theories may profit from a non-linear approach.

Dynamic systems models are not chosen on arbitrary grounds, but because they are convenient in four ways. First, empirical data can be fitted to decide what sort of change one is dealing with. That is, a fit with a growth model could help explain the shape of development. Fits of developmental curves can be a first help in deciding whether or not change is non-linear. Second, these fits have to be explained. That is, why is non-linear behaviour present in development and especially, what are the mechanisms responsible for this sort of change? Third, with the aid of non-linear equations and theoretical considerations a model can be made to describe development. But this description is more than a description if the model is correct, because, fourth, predictions can be made on the basis of the model. So, these models have descriptive, predictive and explanatory adequacy.

Analyses of development are based on theories or on quantitative behaviour. This quantitative behaviour in development (closely connected to the growth curves of Sternberg & Okagaki, 1982 and to the analyses of Fischer, 1983a) has a number of characteristics. First, variables change over time. Irrespective of their increase or decrease (or even stay on the same level), there is a time index (in graphs, this is usually the x-axis). Second, the variables have a score over time. That is, a variable can be measured (the score is on the y-axis). Third, from these quantifications, qualitative derivations can be made to discuss whether discontinuity is present in development. The crucial aspect is that models are quantified.
The definition of development I use is “a score of a variable changing over time”. At any moment in time, there is a score (e.g. the outcome of an observation, or a test score). This score either increases, decreases or remains constant. These scores result in a growth curve (i.e. a collection of points connected with lines). Traditionally, this collection of points (or the resulting growth curve) is explained and described by linear models.

**Linear models**

In psychology, linear models have been used to explain the underlying trend in a series of data. Most important aspect of linear models is that they are continuous and that the independent variable is additive. Here, I confine myself to models of change that are time dependent. Thus, the time effect of the next point is added to the present point. In other words, there is linear relationship between the independent (time) and dependent variable(s).

The classical linear regression model, i.e. an equation for a linear relationship between the dependent and independent variable, is in (3.1).

\[(3.1) \quad L_t = a + b \cdot t\]

An additive effect means that the effect of time on the course of a dependent variable is at any time equally high.

\[(3.2) \quad L_{t+n} = a + b \cdot (t + n)\]

The equation means, in words, that the level of a dependent variable is the sum of a constant $a$, and a term $b$ (the growth ratio) times any time index $t$ (e.g. age). The equation is additive in the sense that for every value of $n$, the time effect is additive. The dependent variable is dependent on time, not on its own previous value.

Linear equations do not necessarily take the form of a line, since polynomial equations are also linear. The generalised quadratic linear model, which allows fits of curvilinear changes, is equation 3.3. Like in equation 3.2, equation 3.4 shows a linear effect for every $n$.

\[(3.3) \quad L_t = a + b \cdot t + c \cdot t^2\]

\[(3.4) \quad L_{t+n} = a + b \cdot (t + n) + c \cdot (t + n)^2\]
New terms (e.g. \( d \cdot (t + n)^3 \)) may be added to the equation, but the relation is still linear, although it is a more complex polynomial equation, in which time is the factor where development is depending on. In the last five years or so, additional continuous models have been developed to explain development. These new models are non-linear.

**Non-linear, continuous models**

*A dynamic systems model of cognitive development*

A number of studies assume that non-linear trends exist in development. These trends have been called non-linear because the trends of growth curves are, with respect to their growth over time, either gradual but not linear (cf. Dromi, 1986 on the development of vocabulary) or they occur suddenly (cf. Brown, 1973 on the development of function words). Non-linearity means that age or time in general cannot be the independent or explanatory factor. Factors that determine non-linearity are treated in detail in this section.

The non-linear theory that is discussed as an example is a growth theory (van Geert, 1991). The growth model proved to be an adequate model of development. With the aid of an example from the literature (Wijnen, 1996), I discuss the conceptual elements of the model (i.e. growth rate, resources, carrying capacity), their relationship with psycholinguistic entities, and a mathematical model of behaviour.

It is important to notice that any model is a simplification of the real human behaviour. In this chapter, this simplified behaviour (i.e. a model) is, however, not just all behaviour of all children (for example, averaged over age). The model is build on the basis of *individual paths of development*. Furthermore, any model can be based either on theory or on data or on both. Only a theory or data is presumably not enough to fully comprehend the nature of development. Using data only could lead to a nice fitting result, but any highly complex model could be used to fit data perfectly. A model of non-linear equations that is based on a theory only might have fitting results that do not follow the path of development.

An extensive treatment of how to apply dynamic systems modelling to cognitive and language development can be found in van Geert (1991). To explain the parameters in the model and also how to build a model of (language) development, I use the following example throughout this section. Suppose that a child is learning a rule (in this case a rule about language), namely the rule ‘The inflection of verbs. The total amount of knowledge (i.e. the knowledge of all situations where this rule applies) can be expressed as a 100 \% score on a test. It is debatable whether adults will use this knowledge in all
situations correctly, but suppose this is approximately the case. The process of arriving at this 100% score, the change, might be linear or gradual, or even a sudden jump. The mathematical equations to model this process of inflection consist of several parameters: a starting value, a growth coefficient, a final state, and (possibly) supportive and competitive factors (since it is assumed that processes are not independent from sources or other variables or processes).

One of the main concerns of this study is how to account for quantitative growth in development and how to relate this quantitative growth to qualitative analyses and to the parameters in the model. The definition of such quantitative growth (or development) is “an autocatalytic quantitative increase in a growth variable following the emergence of a specific structural possibility in the cognitive system.” (van Geert, 1991, 2). In other words, language development (or cognitive growth in general) consists of growth which is expressed in numbers (quantitative) and which is not entirely caused by some external factor (autocatalytic). This growth is caused by the system itself, i.e. change is induced by the cognitive system itself. In the case of finite verbs, the growth of finite verbs is, for instance, not induced by the development of verbs in the input (the parents), but it is due to the structure of the system (i.e. language) itself.

There are a number of assumptions in van Geert’s mathematical model. First, it is assumed that language growth is constrained by limited resources. Memory, attention and motivation (among others) are limited during development. To use the example of verb inflection, the speed of learning to use finite verbs is restricted due to these resources. Second, there is an end state of the language system (in the context of this thesis prescribed by a syntactic theory) which is a collection of supportive and conditional factors. These factors result in the carrying capacity, which is in our case, the knowledge of the application of verb inflection that can be achieved by a language learner who is subject to certain limitations of learning and acquisition. Third, growth is hampered by what is called feedback delay. Feedback delay is the effect of a specific growth level on a later growth state. If a child learns to inflect certain verbs, for example, in Dutch so called regular (or ‘strong’) verbs like in (imaginary) example 3.1.b.

3.1.a Oom Willem bakte zoekte broodjes.
‘Uncle Bill baked little sweet bread.’

A child that has learned the inflection rule (in the example: bak + te) will benefit from this knowledge on later growth states. It may even lead to regular inflected irregular verbs like in (imaginary) example 3.1.b.
3.1.b *Tante Hillary loopte de keuken uit. (* means ungrammatical)

*Aunt Hillary walked out of the kitchen.’

Fourth, there are supportive, competitive and/or conditional relationships between growers (since language variables use the same limited resources). Supportive relationships are, for example, the inflection rules for nouns, competition may be found in the two forms of verbs (regular and irregular; see also example 3.2) and a condition for the inflection of verbs is the recognition of verbs is being verbs. A mathematical model of cognitive growth has been built, using empirical data of Dromi (1986) (van Geert, 1991).

In the next section I shall explain the transformation from theoretical or empirical considerations to a mathematical model, based on equations.

Non-linear equations

One of the key concepts of growth models is the end state. The use of a rule (e.g. inflection of the verb) is at the end state, ideally, a 100 % score. This means that the full rule application is an equilibrium. The level parameter $x_t$, which is in our case the relative number of correct inflected verbs at time $t$, is growing from one equilibrium (the initial state) to the carrying capacity (the end state). It is also a time dependent variable (the index $t$), that is, all growers in a growth model are time series variables (e.g. age). I assume that development is an iterative process. That is, present behaviour is depending on previous behaviour. The process depends also on a growth rate $r$. That is, the speed of change of the process is depending on a parameter $r$ which influences the variable $x_t$. In the linear equation in 3.1 this parameter is $b$. The growth rate may be a fixed integer (which is individually defined, and therefore not depending on other factors), but this growth rate may also be a composition of several factors, like memory or other growers. The last option is more realistic, since we assume that no process stands on its own. The various forms of change (linear, gradual or sudden) are expressed by a power parameter $p$. In van Geert (1995), it is explained that the power parameter $p$ specifies to what extent $L$ affects the further growth of $L$. If $p = 0$, equation 3.5 describes an restrictive process, whereas for $p = 1$ a logistic and for $p = 2$ a sudden process is described. Since $p$ may be any real number bigger than zero, an unlimited variety of forms can be described by the equation.

The equation so far will lead to exponential growth (with $r > 1$ and $p > 1$). Of course, this is not in line with normal development which is limited by constraints (e.g. memory). There is an end state in language development. This end state is the body of grammatical knowledge, and it can be built into the equation. In the model, the end state is called the carrying capacity, $K$. The capacity expresses the more or less stable situation: it is not a
constant equilibrium, since it might fluctuate due to random influences. The set of principles discussed so far can be transformed into the following equation:

\[ L_{t+1} = L_t + L_t^p \ast (r - r \ast L_t / K) \]

for \( p \geq 0, t = 0, 1, 2, ..., n, \) and \( L_{t=0} > 0. \)

One of the strong advantages of these models is that relevant psychological entities are embodied in these equations\(^2\), for example, the dependence of behaviour on its previous state(s) and the dependence on external or internal resources. But there is also a possibility of adding competition and support to these models. This competition is any other grower that influences the grower of study. Competition is when two variables in the same system (language) use the same (e.g. psychological) resources in the same temporal space. Suppose that the plural morphology of nouns calls on the same memory units as the inflection of verb. If all attention is used for inflecting verbs, nouns have to ‘wait’ before they get pluralized. This competition can also be found in problems of sentence planning. It seems that when children make longer and more complex sentences, they start to have difficulties in planning the sentence. The problems of planning difficulties result in developmental stuttering (Wijnen, 1990). The competition between attention, memory, etc. on the one hand, and the execution of linguistic rules on the other hand may be added to the model as a factor \( C_r \). Models of competition and support between several growers are called models of \textit{connected growers}. An example of connected growers is the development of multi word sentences. If a child has no words, he cannot make sentences. Words are building blocks of sentences. Thus the development of sentences is connected with the growth of vocabulary. This modelling of connected growers, i.e. a sequence of stages, on the basis of dynamical systems theory has been applied to solving cognitive problems (Fischer & Granott, 1995) and to the development of knowledge of children (Case, 1992).

An important question is whether \( K, r, p, \) and \( x \) are changing over time. Should it be assumed that all parameter values are constant over time, or does, for example, the growth rate \( r \) change during development, due to the increase of memory capacity? It seems that the latter is more likely (because of an increase in memory capacities, or

---

\(^2\) There are two sorts of non-linear equations. \textit{Differential equations} are used when the limit of the time interval is zero, i.e. there are no intervals between two points of time. \textit{Difference equations} are used when there is a discrete time interval. Research on development is discrete by definition, since there is no possibility of measuring every moment, although the processes may be continuous (unless they are the result of discrete learning events).
teaching and learning). An increase of K (i.e. a new equilibrium) implies that the system matures. For example, the child has somehow deduced the rule for verb inflection. The knowledge of this rule leads to a new equilibrium. In a model of development, it is possible to use either constant or variable parameter values. The choice depends on theoretical or empirical considerations.

Everything that grows needs a ‘seed’. In a growth model of language development this seed, needed for the language system to start functioning, is the smallest unit in a collection of sounds, syllables, words, sentences, etc. In our model (Ruhland, Wijnen & van Geert, 1995), words are assumed to be the smallest unit. The seed in our model is a word, which is represented by an equation containing a small number slightly bigger than 0.

Various assumptions of feedback delay are needed since actions in the past may not have direct consequences for the presence. This “credit card effect” called feedback delay also needs to be build in, too, since this delay keeps the system from increasing exponentially (in accordance with empirical evidence; see for example figure 4.1 in the next chapter). Feedback delay is defined as the time interval between a growth level and its cause (van Geert, 1991).

Sometimes, a child undergoes a temporary decrease in his performance. This is called regression. The score of a variable decreases and the system quite often loses its stability (i.e. the system becomes temporarily unstable). However, regression does not need to be implemented in the equation as a separate, explicit parameter. Consider a system of two connected growers (e.g. two opposing grammatical rules: the use of verb inflection vs. the use of pronouns). When the new grower starts to grow, this new grower (or rule) might lead to a temporary drop in the performance of the first rule, or lead to a complete disappearance of the first grower. For example, when a child learns a new rule (e.g. the inflection of verbs), the child has less time (in terms of memory or attention) to spend on, for example, lengthening his utterances.

Finally, two concepts are used to describe how patterns arise in a system. First, there is selforganisation. Selforganisation is when a pattern arises in a system, but when no specific rules are given that prescribe the pattern (see Kaufman, 1993 for some examples). Second, an attractor is a state of a system where it is forced to without any explicit instructions. In physical terms, the system searches for its energy minimum. In psychological terms, this means that there are states in a system which are easier to reach, either because of the input, or because of the constraints of the system. Selforganisation and attractors may also be the key to get out of the logical problem of language acquisition. These attractors may be, for instance, based on competition (MacWhinney, 1997).
Applications of the dynamic modelling have been applied to cognitive development (van Geert, 1994a, 1995). Similar and equally promising results have been found in other lines of developmental research. Thelen and Ulrich (1991) found evidence for underlying non-linear processes in motor development. Children have walking skills in the first months of life. This behaviour is present from birth, but it may be dimmed by other factors (e.g. too much fat in the child’s legs, or not enough muscle power). In McCune (1992), the dynamic systems view is used for the development of first words, while Fogel (1992) used the dynamic systems approach for the dynamics of movement and communication. In sum, the dynamic systems approach is useful for all research which aims at describing and explaining processes that do not follow a smooth path along the time continuum.

Non-linear equations form a new collection of models of psychological processes. The purpose of these equations is not just a theoretical exercise, but they are used to fit data and build models. The goal of fitting is to find evidence for a certain kind of change, and to give an interpretation of the equations, and the processes they fit on to.

Dynamic systems equations have three applications. First, there is the possibility of curve fitting. Fitting data means that parameters of equations are estimated to determine the basic form of the development of a grower. In this sense, non-linear models have a statistical purpose. Most of the time, data are tested against assumptions of linear regression and unimodal distribution (see the next section on other than unimodal distributions), but data can be tested against non-linear regression models.

Second, empirical curves can be predicted. There are two methods for predicting empirical data. So-called genetic algorithms take as an input a few data points and calculate the next (unknown) data points\(^3\). This method is extremely useful for, for instance, prognoses on the capital market. Second, predictions may be based on connections between growers. This means that if the relation between two growers is known (i.e. extracted from theoretical considerations), this relationship can be modelled with non-linear equations. Running simulations with these connected growers shows how a change in the development of one grower leads to change in the other. For example, if for theoretical or logical reasons multi word sentences (sentences longer than one word) can only arise if vocabulary has developed satisfactorily (i.e. vocabulary is over some threshold), then the grower “vocabulary” is a precursor of multi word sentences. The precursor of vocabulary, for example, is dependent on phonological and phonetic structures and processes. In sum, every variable has a precursor that is a

\(^{3}\textit{Neuralyst} \text{ is a program that uses genetic algorithms. However, it would lead to far to go in full detail in the theory of genetic algorithms in the context of this book.} \)
summary and reduction of other systems or partial subsystems (e.g. memory, neurons, attention, etc.). In the predictions of the change of growers, this reduction is necessary to avoid an infinite regression of precursors (with respect to the precursor of vocabulary, see McCune (1992) for a study on the dynamics of first words).

Third, equations are used to build models of development. These models can explain how certain behaviours arise, but this explanation definitely needs a structural theory which specifies possible structures. In language development, these possibilities are given by a linguistic theory (e.g. by trees). For example, the initial sentence structure (CP) of a child might be built up from an Noun Phrase (NP) and a Verb Phrase (VP) (e.g. a subject and a non-inflected verb). If we take as a start a reduced competence hypothesis (see for example Radford, 1988b; see also chapter 4) which entails that some branches in the linguistic tree are not available from birth (only lexical categories like NP and VP are available), the question is in what fashion new branches are added to the tree. Do they grow instantaneously and all at once? Some hypotheses assume that all functional branches of a linguistic tree are available from birth (a linguistic ‘strong continuity point’ of view), but whatever the assumption, development itself might be hampered or limited by memory or attention. The growth model and the linguistic theory are connected as two complementary issues. A linguistic theory provides a valid argument for the choice of the variables, and the theory also specifies the relationships between these variables. In other words, a linguistic theory specifies the structural terms that are translated into quantifications in a growth model. These linguistic theories also explain to what degree the assumptions of innateness predict the course of development. To recapitulate, a growth model must take into account a structural theory that explains the order of development. The structural theory is also needed as a source for the choice of the variables.

Summary of continuous non-linear models

In sum, growth models are complementary to structural models, i.e. linguistic models. In growth models, a process is split up in various elements that are based on theories of development. Through the assumptions of these theories these elements are translated into parameters that are linked to a psychological reality. These parameters in non-linear equations (the important element in growth models) interact with each other, and describe quantified behaviour (in contrast with structural models that are qualitative in nature). Dynamic systems theory and growth models explain the interaction between a grower and its resources, the connection(s) between growers, and the characteristics of a process. Linguistic models and theories are needed to choose the variables that express
the development of language structures, and these models explain the order, i.e. the way words, constituents (e.g. ‘in the pub’ in the sentence ‘I am drinking a couple of pints in the pub’) and sentences are acquired by a child, of the acquisition of language structures in development.

The models that I presented are continuous, but non-linear models. These models exhibit strong accelerations, regressions, and fluctuations. Growth models are also very useful for modelling relationships between several variables. However, these models are continuous models, so discontinuous behaviour cannot be modelled. That is the reason to introduce Catastrophe Theory in the next section.

### 3.2 Evidence for discontinuities in processes: Catastrophe theory

**Introduction: What is Catastrophe Theory?**

Brown (1973) and others found that function words (see chapter 4 for an extensive treatment of these words) appear suddenly in the language of children. The problem is how to define *suddenly*. The concept of suddenness in the development of, for instance, function words remains vague, since it often entails the notion of discontinuity. To overcome the vagueness, discontinuous models are needed. The models discussed in section 3.1 can be applied to processes that are non-linear. Van Geert (1991) has elaborately shown that his models are adequate for the description of cognitive development (e.g. language development). However, these models cannot be used for the identification of transitions (sudden changes) between states: these models cannot discriminate between accelerations and true discontinuities. A subset of non-linear models that can be of help for the explanation and description of transitions (sudden changes) is a theory that offers criteria for transitions and states. This theory is called Catastrophe Theory (CT) (Thom, 1975). CT is a mathematical theory of discontinuous equilibrium changes (or phase transitions). This theory provides empirical criteria (called flags) for the detection of phase transitions. It also provides a formal definition of stage wise development. Since the stage criteria of Piaget have met with substantial criticism, CT can be used instead in order to provide a formal model of developmental transitions and stages. This theory offers criteria that have to be linked with empirical criteria. Note that the definition of discontinuity is the psychological definition, not the linguistic one (see chapter 2). The graph belonging to this form of discontinuity is Figure 2.2 (the right figure).

In short, CT is a theory that explains the characteristics of a phase transition in a system. During a transition, the system is unstable and therefore the system shows
features that are an expression of that instability. Furthermore, the system (language) is measured on the basis of a dependent variable. The phase transition (e.g. expressed by a temporal instability) of this dependent variable (e.g. the rule of verb inflection) is controlled by one or more independent variables. Time is not a control variable in CT. The use of control variables as a way of finding discontinuities in development is called *modelling*. It is the direct specification of behavioural and control variables in one of the elementary catastrophes.

In the next section, I explain why an analysis based on control variables is, in the study of discontinuities of language development, difficult, and maybe even impossible.

*Control variables*

The application of equations (the so-called catastrophe *analysis*) to (cognitive) development is problematic (and maybe even impossible). This has partly to do with the unknown status of the control variables. A control variable is the independent variable of a psychological phenomenon. The behavioural variable is the variable which is measured on a scale or test.

Control variables are needed to observe some of the catastrophe flags. These variables cause the observed change in a system (e.g. language). However, not every independent variable can serve as a control variable. For example, age (or time) increases continually, but age is not a good control variable, because it has been known for a long time that children have individually timed paths of development. Both speed and timing of development vary between children. Van der Maas (1993) has argued that in cognitive development control variables are hard to identify. A possible candidate for a cognitive control variable is memory span which increases over time. In Wimmers, Savelbergh, Beek and Hopkins (1997), several control variables have been chosen to serve as a guide for motor development. For example, for the development of reaching, muscles, bones, and fat may serve as control variables.

In language development, linguistic variables might also act as a control variables. For example, for the development of multi word sentences, words are needed. So, before a child can lengthen its utterances (or make utterances at all), the child must build up its vocabulary. An increase in vocabulary is a continuous control variable for a potentially discontinuous change in multiword sentences. In the end, an analysis of catastrophes must contain an idea of which control variables guide language development through a discontinuous phase. This leads to an understanding how development is guided by factors that cause transitional change, the same way motor development is controlled by
Chapter 3

-50-

the development of muscles, fat and bones (amongst other factors). See also the appendix for a discussion on equations of control variables.

This thesis on change in language development uses a Catastrophe Theory detection. Since I will apply CT to language development, with the stress on syntactic properties of language, the first question is: how can we apply Thom’s theory to (language) development?

**Catastrophe detection**

Apart from catastrophe modelling (see above) and analysis (i.e. a mathematical analysis of the dynamic equations of a transition processes; this method requires knowledge of the mathematical equations), only detection is useful for cognitive development since we do not know what the equations or control variables are that govern a transition (van der Maas, 1993). Detection of a catastrophe happens by using properties or flags that indicate the presence of a transition or catastrophe. The simplest catastrophe model is the cusp and the eight flags.

**The cusp and its eight flags**

Since the empirically observed data show the level of a variable, not its equilibrium level, a problem occurs with distinguishing a mere increase in growth rate from a sudden jump between two equilibrium modes. Catastrophe theory provides an answer to this problem, by specifying a set of eight empirical criteria that are typical of sudden catastrophic equilibrium jumps. All these indicators, also called catastrophe flags, together mark a transition in a developmental variable.

There are 7 elementary catastrophes (Thom, 1975), which differ in the amount of control variables and the power of the equation (i.e. the more control variables there are, the higher the power is). All catastrophes are sudden changes (i.e. discontinuities) which are caused by changes in control variables. Out of the seven catastrophes, one is frequently applied to cognitive development, namely the cusp. This theory increases the knowledge of processes independent of the field of psychology, and in all, the use of Catastrophe Theory is not limited to the study of (cognitive) development. In social psychology, Tesser and Achee (1994) have studied social interactions using Catastrophe Theory. This is also one of the charming benefits of Catastrophe Theory: it is a formal model that may be applied to any process.

The question arises: why use the cusp, and not one of the other 6 catastrophes? There are three reasons. First, most transitions are controlled by changes in a small number of
variables. The cusp demands for 2 control variables. The number of attractor states in the cusp is 2, exactly the number of equilibrium states in the development of functional categories (see next chapter), and it also the catastrophe model with the smallest number of control variables that is needed to explain discontinuities. It is less complex than any other model for discontinuous change. Second, the cusp has proven to be applicable in various areas of psychology (van der Maas, 1993, Wimmers, 1996, Tesser & Achee, 1994). Third, this is only the start of a new area of research using CT. Therefore the simplest model is used.

The cusp looks like a folded table cloth (see figure 3.1) with one behavioural variable ($z$) and two control variables ($x$ and $y$). Roughly speaking, we consider equilibria, that is, stable periods in development, and change between these equilibria as the crucial elements of development. The change between these equilibria can take several forms, but in order to decide whether this change is transitional or not, criteria from CT are applied to language development. Following Gilmore (1981) and van der Maas (1993), I relate discontinuity criteria, i.e. the so-called flags, to developmental research. Van der Maas (1993) has found several examples in the literature where bimodal distributions are present. Although such quantitative findings are lacking in language development, it is generally accepted that during the first months, a child changes from telegraphic speech to differential speech (Brown, 1973). These two stable states of speech are a linguistic form of bimodality. It relates to the first flags in CT bimodality ($I_a$ and $I_b$) which refers to the fact that there are two states (equilibria) in a process. In fact, the bimodality flag

![Figure 3.1. The cusp](image-url)
indicates the probability distribution of the scores in the transition area. The chance in this area that there are not one but two modes is very high.

Van der Maas (1993) has applied this criterion to conservation development. He found a non-conservation stage and a conservation stage. Proposals that there are even more than two states (e.g. a trimodal distribution) are regarded as inappropriate, since, for example, partial conservers are children that are in the middle of a transitional process. Characteristic of partial conservers is a non-constant score on a test during the transition period: one day their score is high (i.e. most answers are correct), the other day many mistakes are made. One must be careful, however, in using this bimodality criterion. Suppose the growth of a test score looks like figure 3.2.1. If the scores are plotted in a frequency histogram, the result would look like figure 3.2.2. Thus, the bimodality flag is not sufficient in order to decide what sort of change is present in development.

Besides that, there is second problem. Namely, how can one test bimodality? Most if not all the widely used statistical procedures assume unimodal distribution. Hartelman (1996) describes a procedure for a statistical test of bimodality. In the transition area, the data are bimodally distributed. Outside the transition area (i.e. the cusp), the data are unimodally distributed.
The second flag helps us a bit further. It is assumed that there is a region that is inaccessible (II): behavioural variables cannot be in a state somewhere in between these two possible behaviours. This flag prevents the conclusion that a gradual change, in which case there are also two states, is a case of qualitative change. If the data were truly bimodally distributed, the distribution would look like figure 3.3.1 and the frequency distribution would look like figure 3.3.2. However, a problem looms because due to error variance scores may appear in the ‘forbidden’ inaccessible area. For this reason, inherent to psychological research, it is accepted that there is a probability that scores fall in the inaccessible region (see also Wimmers, 1996). In short, error variance prevents true inaccessibility.

Inherent to the first two flags is a characteristic of change which has been mentioned, although indirectly, by Brown and others. The transition from telegraphic speech to differentiated speech (see chapter 4) has always been referred to as a sudden change. Suppose that change from one state to the other takes place very slowly, would it be possible to decide where one state ends and the next starts? The answer must be no, since there are no sharp boundaries between the states. The only point that can be made is that at the beginning there is a different situation from the one at the end, but this is extremely vague. Apparently the speed of change, in combination with bimodal distributions and inaccessible regions, is important, and although time is not incorporated in CT, the third flag is dedicated to the speed of change. It states that change must occur suddenly. A sudden jump (III) refers to the fact that a change in equilibrium is abrupt.
A variable jumps suddenly to the second (higher) equilibrium. This is possible at two moments: along the two paths b1 and b2 (see figure 3.1).

The next three flags are indicators of instabilities in the region of a catastrophe. **Divergence of linear response** arises when a system loses its stability after it is perturbed. Perturbations of the control variable(s) near a catastrophe point will lead to a loss of stability, and to large oscillations of the behaviourable variable. Empirically, one needs dense time series (i.e. time series that covers the developmental path of a variable in such a way that every small deviation can be observed) to observe divergence of linear response. The problem in research of cognitive development, however, is that these time series are rare, and the control variable is often unknown. This flag explains why training or corrections in language development seldom leads to success. It has been reported, more than once, that if a parent corrects ungrammatical utterances of a child, the child hardly ever pays attention to these parental corrections. Here follows an example from Braine (1971) (cited in Ingram, 1989):

Child: Want other one spoon, Daddy.
Father: You mean, you want THE OTHER SPOON.
Child: Yes, I want other one spoon, please, Daddy.
Father: Can you say ‘the other spoon’?
Child: Other ... one ... spoon.
Father: Say ... ‘other’.
Child: Other.
Father: Spoon.
Child: Spoon
Father: Other ... spoon.
Child: Other ... spoon. Now give me the other one spoon.

3.1 An empirical example of linguistic stability.

The example from Braine (1971) is not a real example of divergence of linear response, since an empirical example of this flag in language development does not exist. This flag calls for learning or imitation, and that is what is neither expected nor hypothesised in language development. What the example illustrates is that CT predicts that only near a catastrophe point these corrections may have an effect. Likewise, training before and after the transition is very likely to be ineffective. In both cases (before and after a transition), the equilibria are too stable to be perturbed and eventually change to a higher level.

**Critical slowing down** refers to the fact that a system that is perturbed near a catastrophe point needs more time to get back to its (old) equilibrium. This means that
its relaxation time will be longer. This does not mean, however, that the quality of change differs when a system is perturbed: only the time a system needs to get back to its equilibrium is referred to.

This flag is not related to any empirical findings in developmental psychology. However, it is possible to invent an experiment in which a child must repeat sentences in a training session. Suppose a child is in his telegraphic stage (see also chapter 4). One prediction would be that the child will have longer reaction times (eventually leading to non-responses) to the more complex and difficult sentences. Transitional children, i.e. those children that leave telegraphic speech but who have yet to reach differential speech, will show slowed down reaction times, whereas children who have reached differentiated speech, master all sentences without hardly any hesitation.

Anomalous variance is when behavioural variables show an increased variability near a transition. As van der Maas (1993) observed, this has two consequences. First, correlation values will drop near a catastrophe. This means that the structure of the system changes. Old coherences disappear, whilst others have yet to appear. Second, the new equilibrium (i.e. after the jump) may show considerable oscillations that die out when a variable gets further away from the catastrophe. This means that if a child is capable of using a new syntactic rule, the dependent variable might fluctuate during some time around the point of transition.

Like critical slowing down, there is no clear link with existing empirical findings. However, this flag refers to the fluctuations during the transition period which may be found in the score on a test. A second, possible indication that a dependent variable is subject to anomalous variance is when children show inconsistencies in their verbal reports (van der Maas, 1993), or discrepancies between their verbal and non-verbal behaviour. These discrepancies between speech and non-verbal behaviour have been found in children who show in their non-verbal behaviour the opposite behaviour of their language (e.g. acting out experiments; van der Wal, 1996, 108/109). Goldin-Meadow, Alibali and Church (1993) also found a gesture-speech mismatch. They explain this phenomenon as a transitional phase, in which children show a higher error variability in both the cognitive and motor domain. This increased variability points at a loss of structural stability.

The last two flags are vital in deciding the true nature of a catastrophe. The first flag is divergence. It occurs when a small change in the initial value of a path through the control plane ultimately leads to large changes in the behavioural variable. In the cusp (figure 3.1), divergence occurs when the control variable changes along the y-axis (from the back to the front) and the behavioural variable has to follow one path or the other. This leads to either the top of the cusp (one stable solution) or to the bottom (the other
stable solution). This is what van der Maas (1993) has called the splitting value. If this value is small, there will be a small jump only (or just a acceleration), but if it is high, the jump (as a function of the splitting value) will be large. It must be said, though, that experimental conditions have to be optimal, otherwise divergence cannot be observed. If the setting is not optimal, the result on a test will lead to a unimodal distribution. Again, this shows the strong interdependency of the flags. Divergence shows another feature of the cusp, namely the bifurcation set. A bifurcation is literally a “twofork”. This bifurcation set explains why an initial small difference between two paths in the control (i.e. the plane controlled by X and Y) ultimately leads to very large differences (van der Maas, 1993, 14).

Empirically, divergence is closely associated with training effects. The example on the previous page refers to this training effect. As long as the child is not near a point of divergence training will not lead any changes in the language of a child. However, when a child ‘is’ in the control plane, small effects of training may lead to the acquisition of language variables.

The final flag is called hysteresis. This flag involves transitions on different points on the time scale, that is, there is a difference in transition points occurring when a control variable either gradually increases or decreases. This can best be explained by following the two paths of b in figure 3.1. The change along path b1 shows a jump. That is, if a variable along the x-axis is increased (from left to right), the variable has to jump somewhere from the bottom to the top. This jump is ultimately on the right side of the tablecloth. The change along path b2 (from right to left) is on the top of the folded tablecloth, but ultimately it has to “jump” on the left side of the cloth. This discrepancy of path b1 and b2 is called hysteresis. Hysteresis has been observed in visual experiments (Ta’eed, Ta’eed and Wright, 1988).

In development, this flag has not been found, although some experiments on conservation show some vague remeniscence of hysteresis (cf. van der Maas, 1993; Boom, Gerlagh & den Hartog, 1997). With respect to language development, categorical perception of sounds like b and p (in phonological terms; the voiced and voiceless bilabial explosives) is supposed to be relatively abrupt, i.e. with a small fuzzy overlapping area (Eimas, 1985). In a test, when a baby is offered a series of sounds where a b changes into a p or a p changes into a b (i.e. in the overlapping area), the perception itself may lead to different jumps, depending on which sound was presented first. Evidence for a difference in perceptual switching can be found in, for example, Bohn and Flege (1993).

In development, this flag can not be observed in longitudinal data, because the direction of x-axis (for example, time) has to be reversed to go from left to right and vice
versa. Only in an experiment, it is possible to find hysteresis. Suppose a child is offered 10 sentences that increase in complexity. The child has to repeat the experimentator (a so called EI-test; cf. Ruhland, 1991), and the child is capable of imitating some but not all of the sentences. Somewhere the child will fail to imitate (e.g. sentence 7). If the child were to start with the difficult questions first, his failure to imitate would probably last to another sentence (e.g. sentence 3). Although longitudinal data are used in this study and therefore this flag cannot be found, hysteresis could be found with such a test.

An important concept of CT is structural stability. Perturbations of a system under developing constraints will not lead to major reorganisations. The system (i.e. language) develops without any reorganisations or instabilities, except for that point in time where the system undergoes the transition. Structural stability means that empirical findings are similar in repeated conditions and small changes in these conditions do not alter the results dramatically. Another aspect is that time is not explicitly formulated within CT. CT is a local theory and model. Changes may be catastrophic, and still take place over a longer period. For example, Zeeman’s dog changes its mood within seconds (Zeeman, 1976), whereas change in children learning the process of conservation takes place over 2 years (van der Maas, 1993). Nevertheless, the jump itself (on the underlying equilibrium level) is instantaneous.

**Evidence for catastrophes in (developmental) psychology**

I have dealt with one of the seven catastrophes, namely the cusp. The eight flags that mark the cusp can be linked to the behavioural phenomena in general and to psychological development in particular. Zeeman (1976) modelled the instantaneous behaviour of dogs when they were brought to distress. There is no continuum: dogs suddenly change from attack to retreat or the other way around. The two equilibria, attack and retreat, have two control variables: fear and rage. There is no intermediary state of behaviour. The dog either attacks, or retreats. The two modes, the lack of an intermediate behaviour, and the suddenness of behavioural change make Catastrophe Theory very suitable for explaining the empirical findings. Other studies have found catastrophic change in psychological processes. Visual perception, i.e. the perception of multistable figures, has been modelled with Catastrophe Theory (Stewart & Peregoy, 1983).

Van der Maas and Molenaar (1992) report on a study of conservation development. Several indications have puzzled researchers over the years. Children do not have the ability to apply conservation before the age of (roughly) 5. Within 2 years, a child learns how to conserve quantity, length, etc. Van der Maas and Molenaar show that this
process is characterised by catastrophic change, since there is a sudden change (sudden jump flag) from a non-conserving to a conserving state (bimodality). Furthermore, they found inaccessibility and anomalous variance. Van der Maas and Molenaar have reformulated Piaget’s equilibrium theory into a formal mathematical theory. Piagetian problems like the lack of a stage and a transition definition have been solved.

In the field of motor development, Wimmers (1996) has performed a study in the development of reaching and grasping, showing that the transition from non-grasping to grasping is a catastrophe-like change. In sum, these studies reveal that the application of CT helps to model and to understand the nature of (non-)developmental processes in psychology.

Catastrophe theory provides clear criteria to determine the presence of discontinuities. There are eight flags, which can be linked to empirical research. CT is a formal (i.e. mathematical) model of phase transitions, that like any other formal model of language development, needs a structural model (see chapters 2 and 4).

3.3 Summary and conclusions of the chapter

One of the fundamental problems in the study of human development that still awaits a solution is the proof of psychological discontinuity (see also chapter 2). In the past, there have been no clear definitions of discontinuity and no clear criteria to demonstrate discontinuity, either. A special branch of mathematics called non-linear models is helpful for fitting and modeling development. These models can describe behaviour processes that show chaotic, non-linear continuous or discontinuous behaviour. That is, behaviour that does not follow smooth paths along the time continuum.

Dynamic systems models are helpful in four ways. First, empirical data can be fitted to decide what sort of change one is dealing with (i.e. a help in deciding whether or not change is non-linear). Second, they incorporate and explain the mechanisms responsible for change. Third, a model can be made to describe development. Fourth, predictions can be made on the basis of the model.

There are two types of models and theories in this branch of mathematics, both capable of explaining several characteristics of development. First, there is a group of models called dynamic systems models. These models are based on non-linear (differential and difference) equations. The models are reductions of individual paths of development (thus, not averages of groups), and it is assumed that cognitive growth is constrained by limited resources (e.g. memory, attention and motivation), that there is an end state of the cognitive system which is called the carrying capacity, that cognitive growth is hampered by what is called feedback delay (i.e the effect of a specific growth
level on a later growth state), and that there are supportive, competitive and/or neutral relationships between growers. Cognitive growth is expressed in a number of parameters in the equations. There is an initial state, $L_0$. $K$ is the carrying capacity, that explains the final state. A variable (or grower) grows from $L_0$ to $K$. The process depends on a growth rate $r$ that controls the speed of change. The exponent parameter $p$ controls the power of the equation. This parameter explains to what extent the current level of a grower specifies the further growth of $L$.

*Catastrophe theory* is a formal theory of discontinuous behaviour. One of the seven elementary catastrophes, the cusp, proved to be very useful for the detection of catastrophes. Eight flags together indicate a catastrophe. Some of them (a sudden jump and bimodality) are not new in research on development. Catastrophes are characterized by instabilities between two modes (or equilibria), a rapid change between these equilibria, and perturbation of the developmental system during the catastrophe leads to either a new equilibrium or to a labourious restoration of the old equilibrium.

The research questions may now be reinterpreted. In chapter 2, I argued that theories on language development lack a thorough formal model for the proof of assumed discontinuities processes in development (i.e. discontinuity in the psychological sense; cf. Piaget’s equilibrium theory). The solution for the lack of models for the study of discontinuous behaviour is a branch of mathematical models that are called non-linear models, i.e. continuous non-linear and discontinuous models. They test the non-linear assumption(s) of development. The study of discontinuities in development has profitted by Catastrophe Theory, which provides us with criteria that can demonstrate discontinuities in development. On the basis of Catastrophe Theory, the central research question now reads:

> Is there evidence for continuous or discontinuous (both in the psychological, i.e. quantitative, meaning) development as defined within the linear and growth model paradigm (continuity) or as defined by the catastrophe flags to language development (discontinuity)?

The next chapter contains the empirical basis of this study, which consists of the children that participated in the study, the linguistic variables, the method of analyses, and the length and frequency of recording and the transcripts.