Rick Ruhland

Closed class words in motion: An application of catastrophe theory to language development

Abstract

In this paper, catastrophe theory is used to decide whether or not the sudden introduction of so called closed class words in the language of a child is a discontinuous change. On the basis of the spontaneous language of six Dutch children, four 'flags' were translated into empirical indicators. Empirical evidence only supports a rapid introduction of function words (i.e. a sudden jump) and multimodality, not a phase transition. This sudden change is accompanied by other features (e.g. by temporal dysfluencies). Explanations about order and form of development are based on linguistic theories, and the relationship between non-linear models and psycholinguistic models needs to be explicated.

Introduction

Change in language development research has traditionally been studied and explained with an emphasis of structural analyses, e.g. from a syntax point of view. That is, theories and models are applied to the grammatical development in terms of (underlying) relationships between words within a sentence. These theories stress the what and why of development, but not the how. Elsewhere I argued that a quantitative, or developmental, approach is needed to fully comprehend the development of language structures (Ruhland, 1997). However, linguistic theories that describe these structures fail to explain the quantitative patterns that exist in development, whatever these patterns are.

The main theme of this paper is not the *acquisition* of these language structures, for instance syntactic rules (i.e. rules that specify relationships between words in a sentence), but the *development* of these syntactic structures as quantitative patterns. I stress these words in italics for two reasons. First, development refers to the meaning of change in a *quantitative* way, where acquisition refers to change in a *qualitative* way. Second, although development (i.e. change) is very important for a model of language development, *development* is underexposed by structural theories and it therefore needs more attention.

The goal is to carry out an empirical analysis of language structures. This analysis emphasises quantitative aspects, and the change found is accounted for

by structural theories, i.e. linguistic theories. These linguistic theories come in two shapes. First, *origin* theories assume that children have linguistic knowledge to a more or lesser extend (with birth). There are several assumptions of competence that try to explain missing categories and word classes. Full competence, for example, means that all rules for the syntax of any language are available to a child from birth on, whereas no competence means that a child has no competence at all. However, these assumptions cannot explain the quantitative patterns in development. Second, theories on *change*, e.g. Parameter Setting, explain the way a child learns relationships between words as intrinsic properties of the syntax of a language, but again, these theories cannot predict nor explain patterns of change explicitly.

In the discussion on human development, the difference between continuity in development has been the heart of many discussions. This difference has proven to be problematic. Over the years, a model of change lacked a formal model to distinguish between the two forms of change, although several researchers claim implicitly or explicitly that some change in (language) development occurs at least rapidly. The main reason to use non-linear theories, e.g. catastrophe theory, is to have formal criteria to decide for or against these implicit or explicit claims of discontinuity in development. That is, with these models it can be decided whether a rapid change is a discontinuous change or not. In other words, non-linear theories are a welcome addition to linear models to highlight the patterns of development. Since linear models fail to describe or explain discontinuous development, Catastrophe Theory has been developed to have indicators (or flags) to analyse sudden change from a discontinuous point of view (see Ruhland, 1997 for an overview of CT and the flags, and the application to language development).

There is a part of language development that clearly shows the problem of distinguishing between continuity and discontinuity. In the third year of life, children start to use, for example, verb inflections. These inflections mark the end of the telegraphic stage (TS), and they introduce the differentiation stage (DS). The change from TS to DS is supposed to be very sudden. These stages have been found in English, but it is unclear whether these stages can be found in Dutch, and whether the change between these stages is a discontinuous change in terms of catastrophe theory.

The question reads whether in Dutch there is a rapid change from telegraphic speech to differentiated speech, and if so, whether this 'transition' is discontinuous. It has been suggested (and proven for the development of English) that some language properties or structures like personal pronouns are acquired very rapidly. However, it is unclear whether this marks a real discontinuity. These properties like pronouns are linked with functional categories, and they are extremely important in language (development). The questions are:

1. What are functional categories, how do they relate to closed class words (i.e. function words), and what are the predictions with respect to development?
2. What are the paths in development (in terms of linearity and continuity) with respect to FC and CCW (or function words)?

Non-linear models (instead of linear models) are applied to quantitative data (i.e. frequency analyses). Structural (i.e. linguistic) theories are used to give reasons for the change found. These theories also explain the importance of functional categories, function words and their characteristics in development.

Functional categories and function words

Early child language consists of one word utterances. In the second and third year of life, children gradually increase the length of the utterances, and the sentences become more complex in structure. Although children do not possess all linguistic capacities after the first 5 years of life (e.g. children acquire some of the semantic structures after the age 5; see for example Ruhland, 1991), there is order in the way structural features of a language are acquired by a child.

Of all language variables, one group of words has been studied quite intensively, namely, the development of *closed class words* (CCW), or function words. There are three reasons why these words are important and why they are an interesting subject of study. First, they are syntactically important, i.e. they constitute the surface form of functional operations in a language (e.g. the inflection of verbs). Second, because of their surface form, they can be measured easily. Third, in the context of the application of catastrophe theory, function words are assumed to be acquired by a child in a very sudden way.

Function words have the following characteristics (see also Gerken, Landau, & Remez 1990).

1. It is a class of words that does not expand over time (hence closed class);
2. Related to functional categories in a language, i.e. associated with syntactic structures;
3. Their distributional properties in sentences are systematic: they appear in fixed positions;
4. They often are monosyllabic (see also 1.);
5. They receive weak stress and may undergo cliticization (i.e. be attached to other words);
6. They are extremely frequent in speech;
7. No dependence on the subject of a conversation.

A few examples of CCW are (in bold):

Articles	1. **The** man gave **a** beggar some money.
Modals	2. I **would** like some whisky during lunch, but I **might** run into trouble.
Personal Pronouns	3. **She** is the professor, **he** is the student.
Prepositions	4. **On** my way back, I found gold **in** the hills.

The problem is, with respect to development, that there is hardly an account for change, i.e. the development of CCW. There is only a rough distinction between

telegraphic and *differentiated* speech. Telegraphic speech is poor with respect to CCW. That is, telegraphic refers to telegram: if one writes a telegram, redundant words are left out, which are mostly likely CCW (e.g. pronouns and articles). Early child language is characterized by a similar absence of close class words.

Brown (1973) noticed that a child abandons telegraphic speech very rapidly by introducing function words, which in turn leads to differentiated speech. This introduction happens between the second and third year of life. So, if the concept of non-linearity (e.g. catastrophe theory) is to be applied to language development, this apparent transition is a good starting point.

Method

Non-linear and linear models: techniques and procedures

Discontinuity is with respect to development the strongest claim that can be made, since more criteria have to be met. Catastrophe Theory offers indicators for the study of discontinuous processes of developmental phenomena. This theory of sudden phase transitions incorporates several features of systems that undergo a change. This change is accompanied by instabilities during the transition from one state to the next. The change between equilibria can take several forms, but in order to decide whether this change is discontinuous, criteria from CT will be applied to language development (see also Ruhland, 1997).

Apart from this discontinuous theory, there are non-linear, but continuous models. Van Geert (1991) developed a growth theory of cognitive and language behaviour of logistic equations. This model of cognition and language is applied to the data of this paper to find out how a non-linear growth model fits the data.

Lastly, linear models are tested against the data. These linear models are in fact regression models that are standard in psychology. In this paper, I used a curvilinear model, i.e. a quadratic linear model.

Design of the study

The analyses are based on the spontaneous observation of the language of six children, starting at 1;9 (92 weeks, Sd = 10.8), ending at 3;1 (159 weeks, Sd = 14). This spontaneous language was collected in sessions of one hour. The children were recorded in a home situation, where both children and parent(s) were engaged in their normal daily routine. In these longitudinal time series, the mother-child dyads were taped every two weeks, and transcribed according to Childes conventions. On the basis of the recordings, frequency analyses of closed class words (e.g. pronouns, modals and articles) were performed with the use of Clan programs of CHILDES (Higginson & MacWhinney, 1991). In the rest of this paper, the focus is on personal pronouns.

Abandoning telegraphic speech: the results

Catastrophe theory: evidence for the flags

As mentioned in the introduction, language variables have been chosen on linguistic grounds and quantified in order to find catastrophe flags and to apply fitting procedures. The first step was to plot individual growth curves by plotting the frequency scores as a function of age (Figure 1).

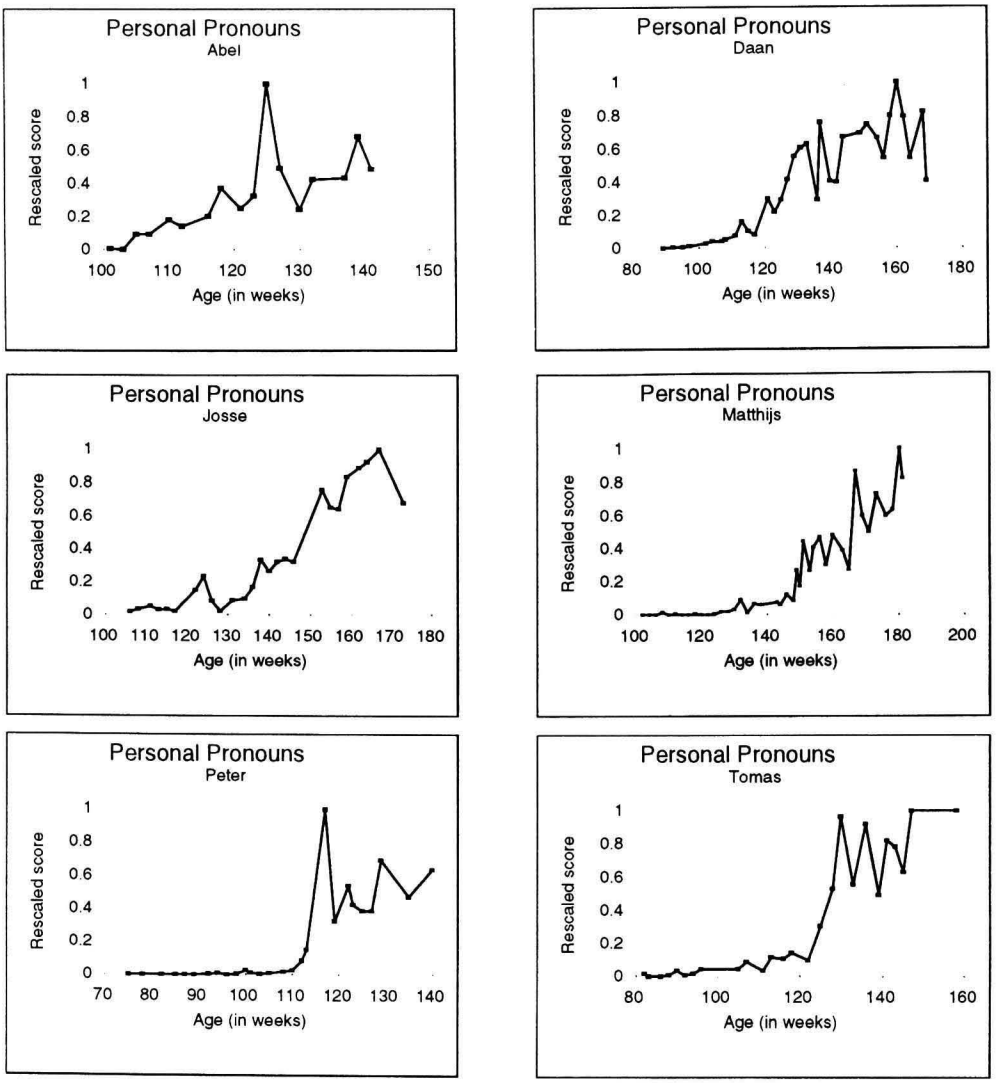

Fig. 1. Individual growth curves.

Through visual inspection, all children but one (namely Abel) showed a rapid increase (from zero to an equilibrium between 0.5 and 1) between 110 and 150 weeks of age. The first flag tested was the sudden jump, one of the indicators of Catastrophe theory. A fit procedure (Ruhland & van Geert, 1998) using a quadratic curvilinear equation and a logistic equation was applied to the growth curves, but before some of the data are presented (see Ruhland, 1997) for data on other CCW), it is important to note that I do not want to start a fitting exercise. The fits are only used to find the patterns in the development of function words. The results were:

Table 1. Fits of the individual growth curves.

Child	Linear quadratic	Non-linear growth
Abel	.71	.78
Daan	.79	.81
Josse	.91	.95
Matthijs	.8	.82
Peter	.73	.88
Tomas	.79	.87
Average	.79	.85

The non-linear growth fit was better than the linear quadratic fit (Wilcoxon signed ranks; $p = .027$). Elsewhere, we argued that high growth rates and high exponents are indicative for a sudden jump (Ruhland & van Geert, 1998, Ruhland, 1997). Therefore, the rapidness of CCW-development was also found in the development of personal pronouns in Dutch.

The second flag, bimodality and inaccessibility were the second and third flag to be tested. In figure 2 the data of the children were plotted in a frequency distribution (in nine steps). It seemed that there was more than one mode.

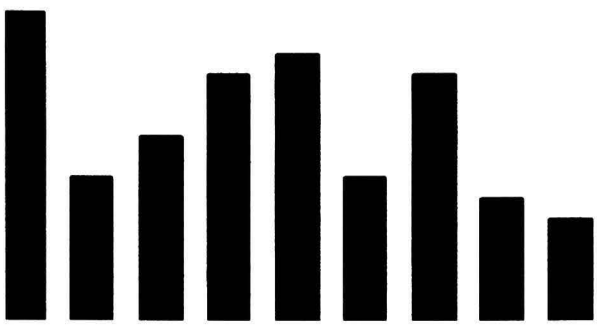

Fig. 2. Frequency distribution of Pronouns.

A testing of bimodality is possible in two ways. A program called Modes calculates the number of modes (or equilibria) in the data of all children on the basis of a kernel technique (cf. Hartelman, 1996). The results are in table 2.

Table 2. Modes analysis of personal pronouns.

Modes	1	2	3
Personal Pronouns	.25	.35	.63

The outcome reads that there was not one mode (or equilibrium) in the data. Second, a mixture distribution analysis (Thomas & Turner, 1991) was used to test the modality in figure 2 (AIC = Akaike's Information Criterion is defined as twice the loglikelihood + the number of parameters in the model, VAF is Variance Accounted For under any model).

Table 3. Mixture distribution analysis.

Model	AIC	chi^2	VAF
1 component	1198	20708,15	0.23
2 components	811	60,1	0.91
3 components	776	2,17	0.99

The conclusion must be that although figure 2 seemed to have a bimodal shape, the statistical analyses did not support this 'visual inspection'. There is no bimodal distribution.

The fourth flag was Anomalous variance. This variance was defined as variance within a session that deviated from a natural linear variance. This anomalous variance was calculated by dividing the second half of the session by the first (Ruhland & van Geert, 1998). This would indicate that if there is a peak in this division, a child uttered more function words in the second half of the session during a transition. Put differently, small problems (like starting problems) mounted up around the transition period.

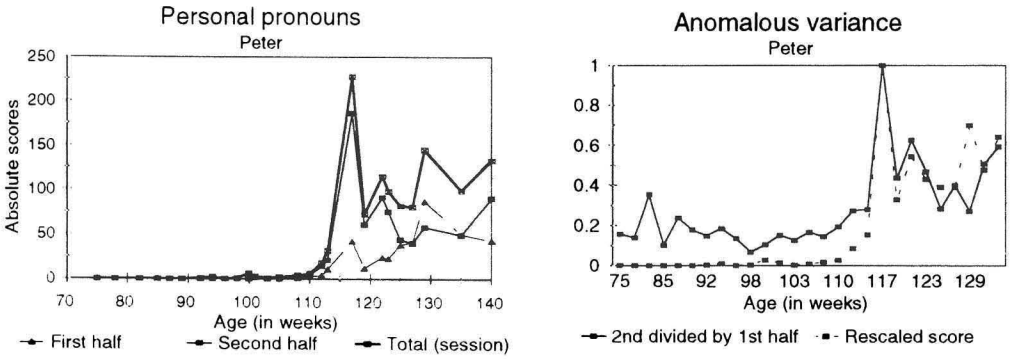

Fig. 3 a and b. Anomalous Variance.

A compensation factor was added to reduce the large influences of small numbers (e.g., compare 2/1 with 200/100). In the case of Peter, it worked out quite well: there was a peak near the sudden jump (see figure 3b). The results, however, were that only two children showed an increase of variance.

Other catastrophe flags cannot be found in longitudinal research. Furthermore, the control variable(s) is (are) unclear, but they are needed to find those flags that indicate instabilities of the system.

In sum, there was evidence that, first, there was a sudden jump in the development of personal pronouns. Second, there was no evidence for bimodality (although there appeared to be multimodality). Third, no inaccessibility could be found, although there was a dip between the modes. Finally, anomalous variance (defined as variance within a session) was found in only two out of six children.

Fits of the data

The fits of the time series are based on a program by Cobb (1981). The program, Cuspfit, fits three models on the data of the six children with respect to personal pronouns. AIC must be the lowest to be the best fit, r^2 is a pseudo r^2 score since there is more than one equilibrium.

Table 4. The analyses with Cuspfit.

Personal Pronouns	AIC	r^2
Linear	637.44	.48
Cusp	289.69	.71
Logistic	391.92	.51

Both AIC and the (pseudo-) r^2 indicate that the cusp is the best fit for the data set of personal pronouns. However, as I argued elsewhere, this does not mean that pronouns change in a catastrophic fashion, since a series of s-shaped curves is also fitted best by a cusp (Ruhland, 1997).

Discussion

It is fair to say that the data do not support a discontinuous pattern in the development of personal pronouns. It is likely that the sudden jump is the result of an acceleration in a continuously developing system. In either case, the discussion has to be directed to the prediction of the patterns found, i.e. continuity and not discontinuity, in relationship with both psychological and linguistic research and theories. For example, if continuity is found in one discipline, the question 'does this mean anything for the other disciplines, and if so, what?' needs to be answered. The explanation of the patterns, i.e. a rapid increase, must have support from linguistic theories on change. Furthermore, it is important that a theory of development incorporates issues of *origin* and of *change* in order to come to a complete theory of development.

I will first summarize other data that are related to the change found, and then discuss the relationship(s) between a non-linear approach and linguistic theories.

Explanations for the change

The findings might be the result of a direct change in two other variables, that explain the development of personal pronouns found in six children. However, I found that although the language productivity of the child increases the personal pronouns do not increase suddenly simply because the children are more talkative (Ruhland, 1997). There is, of course, a relationship between function word development and general language productivity.

Second, the parents are not responsible for the change in personal pronouns. I performed a quadratic regression analysis using several time lags, since a child may not directly follow a (possible) increase in the language of the parents. The results indicate that the parents remain constant in the use of personal pronouns (in terms of the number of personal pronouns per session) in comparison to the use of these pronouns in the children.

The question is: why do personal pronouns increase so suddenly if neither the child's productivity nor the parents language is causing the change? One assumption is that changes in underlying processes could induce a sudden change. This is in fact one of the assumptions of catastrophe theory: a linear or gradual increase in one or more other (control) variable(s), that guide the dependent variable through development. First, one account is that memory increases, either suddenly or gradually. However, there is no evidence for a sudden change in memory (Case, 1985). Second, there is no compelling reason that the development of closed class words is due to an increase of memory, causing them to develop all at once. However, research with connectionist models shows an influence of short term memory as a source for increasing complexity (Elman, 1993). This starting small (in terms of memory capacity) might contribute to the fact that attention is given to only parts of the parent's input. For example, attention is restricted to nouns and verbs in the early stages of language development. In addition, these nouns and verbs have more saliency in terms of non-syntactic reference to the real world.

Related to the increase of memory is evidence for sudden or gradual brain development. As has been pointed out by Bates, Thal & Janowsky (1992), a relationship with grammar exists only on the basis of correlates. Or in their words: 'the relationship between synaptogenesis (i.e. rapid acceleration in number of synapses within and across regions of cortex) and language learning is less direct, i.e. an *enabling relation*' (p. 105). A brain burst provides more storage capacities, which enhances the speed of the grammatical process. This is supported by three facts. First, the process is already on its way. Second, there are also changes in non-linguistic cognition that correlate with grammatical changes. Third, language development is not the consequence of a specific and localized language module that guides the transition from first words to grammar: children can recover from left-hemisphere damage after the second year of life.

Other reasons for the change of CCW in development is found in the stress of closed class words. The evidence is that the omission of function morphemes may be

the consequence of a production constraint, not an encoding limit. That is, the language system is changing because of stress and rhythmic constraints. This has been found in both Dutch (Wijnen, Krikhaar & den Os, 1994) and English (Gerken, Remez & Landau, 1990).

More changes in the developing system

There is more evidence that the system is changing. There is an increase of productivity which is accompanied by dysfluencies in the language of a child (Wijnen, 1990). A child has problems in planning her utterances because there are new syntactical elements that make the utterances longer, but also more complex. The introduction of function words leads to planning problems which are not syntactic. This indicates that several 'modules' in language development (from phonological to semantic, and also non-linguistic modules) are affected by the change in the system, which becomes less stable, or even unstable. Unfortunately, this could not be tested in the language data used for the analyses of the six children in this study, since the transcripts do not have the precision of transcription.

There is apart from an increase of the tokens (the absolute number of personal pronouns) also an increase of types. In other words, the diversity of personal pronouns increases. That is, the average amount of these pronouns just before the jump is 2 types, while after the jump (or transition) the average number of types is 11. This indicates more differentiation in the use of personal pronouns.

Dynamic linguistics

In Ruhland, Wijnen & van Geert (1995) we presented a linguistic reason why function words appear relatively late. We argued that the Case module is triggered by the development of finite verbs (IP). This is in line with, for example, Vainikka (1993). In fact all sorts of functional categories may grow after IP has been acquired. Differentiation of language is build up from a gradual increase in the complexity of the language system. This takes time: the acquisition of a vocabulary, an increase of the length of the utterances, and learning how to inflect a verb leads to the application of syntactical rules using CCW. This increase in complexity could be the result of several processes. We argued, for instance, that maturation might be the underlying process. For the moment it does not matter what the underlying process of structure acquisition is during development. Important to note is that any change, linear or non-linear, continuous or discontinuous, is accounted for by a structural model or theory.

Non-linear models and linguistic theories

Non-linear models cast a different light and shadow on processes in language development since they show that there is definitely something more going on in this developmental stage than a linear change. The majority of the subclasses of closed class words (personal pronouns, for instance) grow suddenly and they all develop

during the same age span (Ruhland, 1997). The analyses in this paper are a useful and necessary addition to syntactic analyses to fully comprehend development: they highlight the process of change.

But most important, I believe, are the two issues that I stated in the first paragraph of this chapter. First, the relation between psycholinguistic models and non-linear models is important because it adds the how to the 'what-and-why' analysis of language development. If continuity is found in one discipline, e.g. in the psychological-mathematical domain, the question 'does this mean anything for linguistics, and if so, what?' needs to be answered. The other way around also applies. If the continuity assumption from linguistics is adopted, then the psychological-mathematical assumption needs to be questioned. Of course, it may be argued whether such a relationship exists, and if it exists, it may be argued what this relationship looks like. For instance, can we map linguistic continuity on mathematical continuity? In other words, quantitative and qualitative patterns need to be linked to support or refute a theory from another discipline. The question is whether an *explanation* of the quantitative patterns, i.e. a rapid increase, must also come from linguistic theories on change. We could refute a linguistic theory that prohibits sudden change in development. For example, if Parameter Setting is correct (i.e. a theory that assumes that linguistic structures appear due to a timing based on language internal structures by setting a parameter which stands for a linguistic fearture), then sudden changes may be possible. One has to be cautious here, though. Structural models and theories are very static, they describe a window in time. The fact that child language is influenced by non-linguistic factors like memory does not play a role in these 'static' theories. Therefore, I think that it is best to keep assumptions of (dis)continuity from different domains separated, since they represent two different kinds of explanations.

Second, it is important that a theory of development incorporates issues of *origin* and of *change* in order to come to a complete theory of development. Traditionally, change was explained with, for example, a linear regression analysis. Instead of linearity, non-linear and discontinuous change can be explained from a formal model like a catastrophe model. In addition to these non-linear models of change, non-linear learning models like connectionism help us out of the origin problem, since all innate assumptions may be abandoned.

We now safely answer the question whether catastrophe theory can be applied to the development of syntactic structures. The answer is yes, but we need new methods, and a different approach to these structures than traditional approaches. Quantifications, even the more rude ones, are by no means a bunch of useless numbers that refer to errors or 'noise'. They are translations of processes that take place. A study of quantified linguistic variables reveals that change is not so 'simple' that it can be caught in linear models. The structural model of language is extended with a process model that is not linear or even continuous.

New method, new insights

Differentiation alone (i.e. a quantitative change) is not enough as an explanation of (language) development. There is no need for new, more complex models if there is

no functional or causal reasoning as an underpinning of the change found. Finding a better fit is nothing more than the application of new methods yielding a higher number (for instance, a higher r^2). These are empty models, since higher numbers do not yield more insight. Likewise, statistics should be used to test empirical findings, to see, figuratively speaking, how bright the light is that shines on the data. Statistics should not put empirics in the dark, and itself in the light. They should be used to establish more certainty about differences between continuous and discontinuous models.

Furthermore, intervention and therapy could benefit from a more intensive study of language development (in terms of the method using dense time series) since these dense time series reveal other developmental changes than linear change. If it is assumed that the non-linear change found in these time series is not unique, then intervention and therapy must take notice of different methods using dense time series that reveal non-linear changes that are the base line of some of language development.

Conclusion

In sum, the data do not convince that the change found is catastrophic. The evidence from catastrophe flags is not conclusive. Only two flags have been found. The use of Cuspfit reveals that the best fit is provided by the Cusp model. However, a collection of rapid 'logistic' curves may have the best fit by a cusp model, since the analysis with Cuspfit only tests the fit of the data in terms of explained variance. Factors like instability of the system are not part of the analyses of Cuspfit. Therefore, more flags must be found using an experimental setting. That setting allows us to test, for example, stability and the sensitiveness on initial values. However, experiments that elicit language production are problematic, especially with young children. A pilot experiment using a Elicit Imitation test did not lead to a satisfactory result. For now, the conclusion must be that closed class words develop rapidly.

If not all functional projections are either present or absent (cf. Ferdinand, 1996), the no competence hypothesis and full competence hypothesis are, on the basis of the quantified data in this paper, at least doubtful hypotheses. A more appealing hypothesis, the reduced competence one, seems to fit the data more adequately. The way reduced competence leads to full competence is that new functional projections are added instantaneously in syntactic development. The behaviour of these functional projections in terms of closed class words or function words follows this 'sudden syntax'. The fits support this assumption, i.e. a sudden introduction of function words, e.g. personal pronouns.

References

Bates, E., D. Thal and J.S. Janowsky (1992). 'Early language development and its neural correlates.' In S.J. Segalowitz and I. Rapin (eds.) *Handbook of Neuropsychology, Vol. 7 (Child Neuropsychology).* Elsevier.

Brown, R. (1973). *A first language: the early stages.* London, Allen and Unwin

Case, R. (1985). *Intellectual development: Birth to Adulthood.* London, Academic Press.
Cobb, L. (1981). 'Parameter estimation for the cusp catastrophe model.' *Behavioral science*, **26**, 75-78.
Elman, J.L. (1993). 'Learning and the development in neural networks: the importance of starting small.' *Cognition* **48** (1), p. 71-99.
Ferdinand, A. (1996). *The development of functional categories: the acquisition of the subject in French.* The Hague: Holland Academic Graphics.
Gerken, L.A., B. Landau and R.E. Remez (1990). 'Function morphemes in young children's speech perception and production.' *Developmental psychology* **26** (2), 204-216.
Hartelman, P.A.I. (1996). 'Cusp Catastrophe Estimation.' Unpublished manuscript.
Higginson, R. and B. Macwhinney (ed.) (1991). *Childes/BIB: An annotated bibliography of child language and language disorders.* Hillsdale, N.J.: Lawrence Erlbaum Associates.
Ingram, D. (1989). *First language acquisition: method, description and explanation.* Cambridge: University Press.
Ruhland, R. (1991). 'De acquisitie van 'hoeven'.' *Tabu* **21** (3), 97-116.
Ruhland, R., F. Wijnen and P. van Geert (1995). 'An exploration into the application of dynamic systems modelling to language acquisition.' In F. Wijnen and M. Verrips (eds.) *Amsterdam Series in Child Language Development: approaches to parameter setting* **4**, 107-134.
Ruhland, R. and P. van Geert (1998). 'Jumping into syntax: transitions in the development of closed class words.' In press. *British journal of developmental psychology*, 16, 65-96.
Ruhland, R. (1997). *Going the distance: a non-linear approach to change in language development.* PhD. thesis.
Ta'eed, L.K., O. Ta'eed and J.E. Wright (1988). 'Determinants involved in the perception of the Necker cube: an application of catastrophe theory.' *Behavioral science* **33**, 97-115.
Thomas, H. and G.F.W. Turner (1991). 'Individual differences and development in water-level task performance.' *Journal of experimental child psychology* **51**, 171-194.
Vainikka, A. (1993). 'Case in the development of English syntax.' *Language acquisition* **3** (3), 257-323.
van Geert, P. (1991). 'A dynamic systems model of cognitive and language growth.' *Psychological review* **98** (1), 3-33.
White, L. (1982). *Grammatical theory and language acquisition.* Dordrecht, Foris.
Wijnen, F. (1990). *On the development of language production mechanisms.* Dissertation Nijmegen.
Wijnen, F., E. Krikhaar and E. den Os (1994). 'The (non)realization of unstressed elements in children's utterances: evidence for a rhythmic constraint.' *Journal of child language* **21**, 59-83.

University of Groningen