Modeling Transient States in Language Change

Gertjan Postma
Meertens Institute Amsterdam

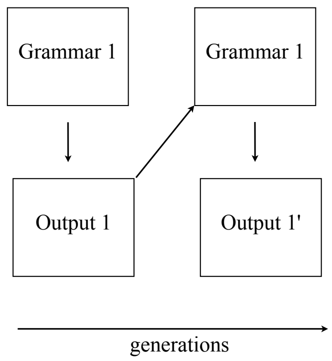
Abstract
Models of language change may include, apart from an initial state and a terminal state, an intermediate state T. Building further on Postma (2010), who observed that the dynamics of the transient state T (failed change) has an algebraic linking of the dynamics of the overall change A → B, we present an generalized algebraic model that includes both the failed change 0 → T → 0 and the successful change A → B. As a preparatory step, we generalize the algebraic function (logist) of two-state change A → B to a differential equation (DE) which represents the law that rules the change. This DE has a bundle of time shifted logistic curves as its solution. This is identified as Kroch's Constant Rate Hypothesis. By modifying this DE, it is possible to describe the dynamics of the entire A → T→ B process, i.e. we have a model that includes both the successful and the failed change. The algebraic link between failed change and successful change (the former is the first derivative of the latter) turns out to be an approximation.

1. Introduction

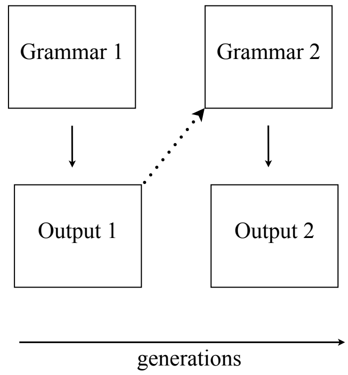
Apart from the standard two-state model of linguistic change, A → B, as discussed in Weinreich *et al*. (1968), Kroch (1989), and others, three-state models has been proposed with an intermediate transient state T: A → T → B (Andersen 1973, Van der Wurff 1990, Weerman 1993). While the two-state model has received algebraic modeling e.g. the logistic model (Bailey 1973, Altmann *et al.* 1983), no overall modeling of the three-state model has been developed. Building further on Postma (2010), who observed that the dynamics of the transient state T ("failed change") has an algebraic linking of the dynamics of the overall change A → B, we present a generalized algebraic model that includes both the failed change and the successful change A → B. Instead of taking a parameterized logistic model as basic, it is proposed that linguistic change is ruled by a differential equation (DE) which has a bundle of time shifted logistic curves as its solution. This is identified as Kroch's Constant Rate Hypothesis. By modifying this DE, it is possible to describe the dynamics of the entire A → T→ B process, i.e. we have a model that includes the successful and the failed change. The algebraic link between failed change and successful change turns out to be an approximation.

2. Three-state models

The existing three-state models all implement the idea that natural language should be split in an I-language or Grammar of L, and an E-language, or set of utterances of L, which are generated by L's Grammar (Chomsky 1986). In stable situations without change, the process of language transmission proceeds along the arrows drawn in the figure in (1).



(1)

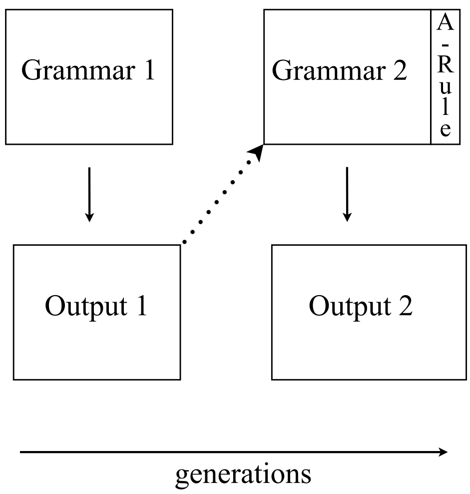


(2)

The recognition that the grammars themselves are not transmitted to the next generation is probably Andersen's. A grammar is acquired via its output. The acquisition process of the second generation is based on the set of utterances that the following generation is exposed to. It is in this step where language acquisition can lead to change, if something distorts the perfect learning, represented in (2), taken from Andersen (1973:767). The imperfection is represented by dotting the arrow. Various theories are proposed about the nature of this imperfect learning, i.e. the nature of the dotted line. We here discuss two types. The first type concerns those theories that situate the imperfection in the learner, i.e. the person that acquired Grammar 2 instead of the target Grammar 1. This type of theory is the oldest and is well expounded in Andersen (1973). The second type of theory situates the change in the adult speakers of the previous generation, who distort their language under sociolinguistic pressure. This is well expounded in Weerman (1993), who is certainly not the first, but who has described it most explicitly. We discuss these below.

2.1 Andersen 1973

Andersen's conception of change situates the change in the new generation that acquires the target grammar by trial and error. Those errors are permanently corrected. Andersen draws his examples from phonology and notes that it is difficult to infer from the perceived *acoustic features* the underlying *articulatory features* in the production. The child of the new generation hypothesizes an (articulatory) grammar and corrects it until it has reached the target Grammar 1. Or not. This happens if the child assumes a grammar that approximates Grammar 1, and makes a tentative correction by an Adaptive Rule (or A-Rule as Andersen coins it), probably interpreted as an Adaptive Rule on the output. This is represented in the diagram in (3).



(3)

The quality of Andersen's proposal is situated in the insight that Grammars are not transmitted directly but acquired through their output. An secondly, and this feature is not represented in the diagram, the fact that Universal Grammar is active in the acquisition process. Thirdly, there is the attractive feature that the change in grammar is situated in the child, as only children before the maturation age can set formal parameters.

Despite these qualities, at least two objections have been raised to this model. The first objection is formulated by Bybee (1982), who noticed that children before the maturation age do not form communities and are, therefore, unable to transmit their language, i.e. the presupposed "imperfect learning" cannot spread. Only adolescent and adult peer groups can. The second objection emerges from the fact that L1 learners are usually seen as perfect learners. Imperfect learning exists but is situated in L2 learning and other adaptive strategies not executed by the Acquisition Device of language but by some other cognitive module (say, the "Problem Solving Device" of Chomsky (1965). This view complies with the metaphor of children as "little inflection machines" (Wexler 1982:44). This criticism is forwarded most consistently by Weerman's observation that "in 'normal' transmission from generation to generation children are simply too good to be responsible for transmission errors." (Weerman 2011:149). This leads us to the second theory of the source of language change.

2.2 Weerman 1993

The other conception of transmission errors situates the source of language change in the adult language use (Van der Wurff 1990, Sankoff & Blondeau 2008), but the most clear advocate is Weerman (1993) who designs a model that we will work out. Weerman tries to reconcile two or three seemingly incompatible ingredients, mentioned in the previous section. These ingredients are given under (4).

(4) - Only children can set linguistic parameters
- Children are perfect learners and cannot be responsible for transmission errors
- Children doe not have the social infrastructure to spread a change

Weerman, therefore, assumes that adult speakers who have acquired a perfect Grammar 1, embellish it during the lifespan with peripheral rules, which are not part of Grammar 1, and presumably not compatible with any parameterization of UG, for instance, because it is a Output Rule that operates in the postsyntax. They do so under peer group pressure, or accommodation in language contact, or by any fashionable linguistic innovation. When exposed to Grammar 1 + Peripheral Rules, the new generation, being perfect learners of UG, (re)set the parameters in order to comply with their parents output and the requirements of UG. As Weerman writes in discussing changes in OV en VO order in the history of English:

> (T)hese (VO) leakages were so to speak exaggerated via L2-M acquisition, both quantitatively and qualitatively. The relevant speakers do not change their internalized setting of the head parameter. From the perspective of their L1 grammar these overgeneralizations are ungrammatical. What they do is add a peripheral rule. A next generation, however, could set the head parameter differently". (Weerman 1993)

This model with peripheral rules is compatible to suggestions as early as Halle (1962), when he writes:

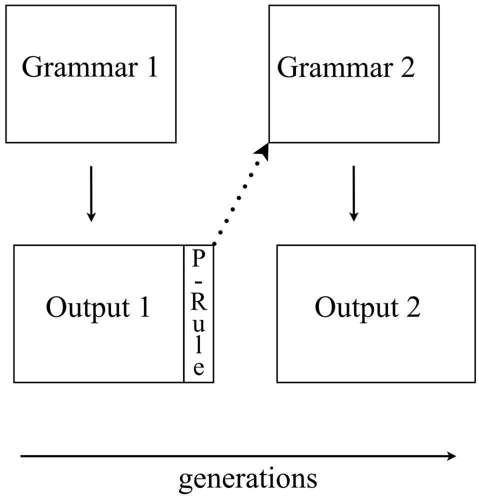
> The language of the adult – and hence also the grammar that he has internalized– need not, however, remain static: it can and does, in fact, change. I conjecture that changes in later life are

> restricted to the addition of a few rules in the grammar and that elimination of rules and hence a wholesale restructuring of his grammar is beyond the capabilities of the average adult.

The idea is also present in Lightfoot (1999:80):

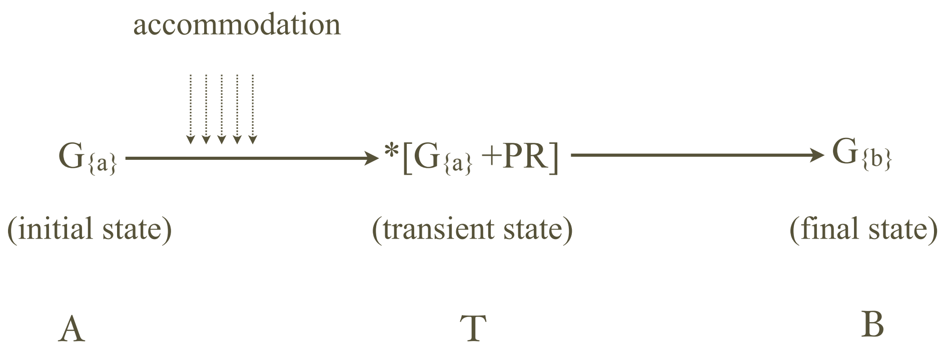
> Although adult innovations may not affect grammars, they (...) reflect changes to the primary linguistic data, the input experience for the next generation of language learners. Adult innovations, then, constitute one reason why an individual might be exposed to PLD which differ from what his mother was exposed to.

This model shares the idea of Adaptive Rule and the idea that the grammatical system is active in language change, that we encountered in Andersen (1973). The two theories, however, differ on the locus of these Peripheral (P-Rule) or Adaptive rules, as temporary deviations from true grammaticality. It is given in (5a).



(5a)

We can describe this model as a three-state model: a system A, consisting of a pure parameter setting of UG, is temporarily embellished with Peripheral Rules (T). This situation which is marked from the perspective of UG, relaxes to a new, pure grammatical state B, produced by a parameter setting of UG.



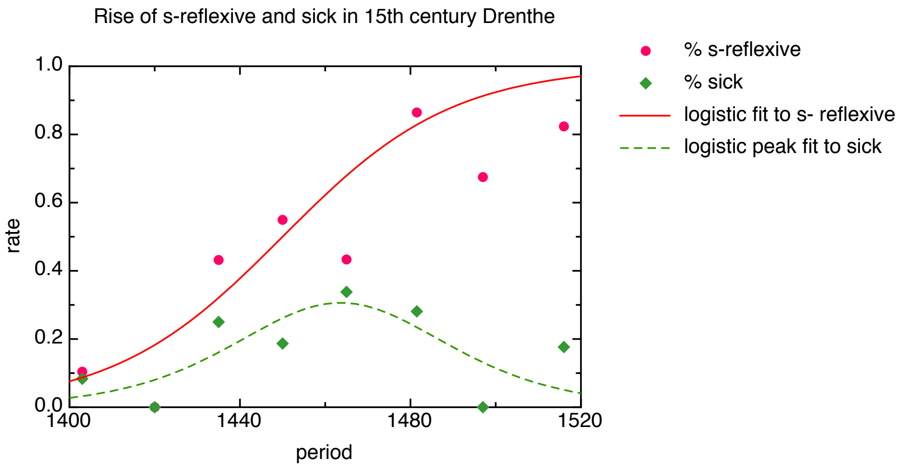
(5b)

The idea is that every change passes through a transient state. A transient state is a state that does not fully comply with the principles of UG. This transient state must be seen as the initiator of the change. In many cases, the Grammar 1 + P-Rule versus Grammar 2 are difficult to separate in the Primary Linguistic Data, but it is assumed that the transient state, being the initiator, is always there. Only in rare cases, for instnace, when another module separates them, they can be distinguished clearly. In the next section, we discuss a case from the literature.

2.3 Illustration

An illustrative instance of this three-state change is found in the emergence of the reflexive in Middle Dutch. Middle Dutch had no reflexive pronoun to express contexts with argumental reflexivity, as in 'John saw himself in the mirror', and simply used the pronoun (as Old English did). In the 15th century, the North Eastern dialects started with the change. The pronoun *hem/om/..* 'him' could not be used anymore, and the language borrowed a reflexive pronoun *sick* from a neighboring dialect. However, this pronoun was not well-formed because of phonological reasons, and was gradually replaced by *sich*, borrowed from another dialect, which completed the change. The curves of the various occurrences are displayed in the figure under (6), taken from Postma (2004). During this century, when *hem/om/...* was gradually replaced by an s-reflexive, we observe a transient period where emerges and fades away.

(6)



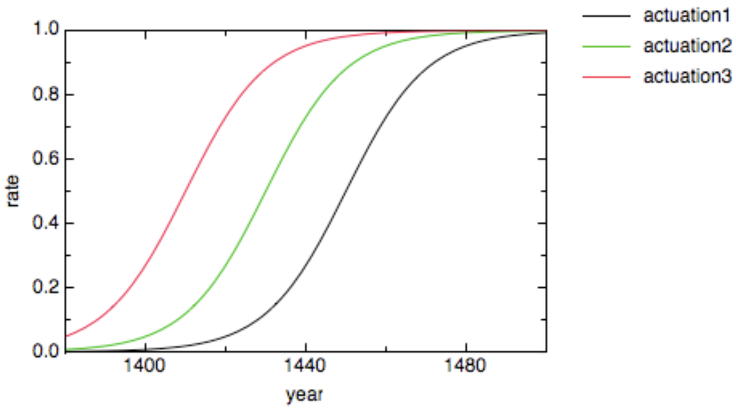
In this paper we propose an algebraic model that captures the basic ingredients of the three-state model. Before we can do so, we need to briefly recapitulate the two-state model: the logistic model with its advantages and disadvantages. This will be done in the next section.

3. The logistic model

The logistic model, proposed in Altmann (1983) and Kroch (1989), is a three-parameter algebraic model that implements the insights in Weinreich *et al.* (1968) quantitatively. While the latter authors take any S-shape curve as the drawn line in (6) as part of their modeling, the former authors specifically opt for the logistic model. The logistic function is given in (7) for further reference.

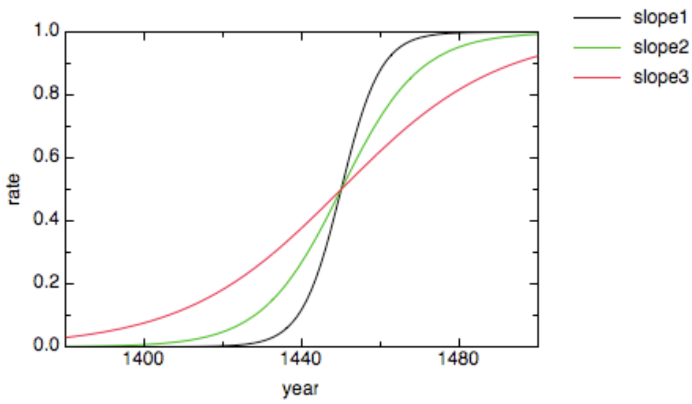
(7)
$$S(t) = \frac{S_{max}}{(1 + e^{-\frac{(t-t_0)}{a}})}$$

Kroch's choice for the logist is motivated by pragmatic reasons. It is a widely used model in the life sciences, and it is easy to manipulate graphically: logistic functions turn into straight lines when plotted on logarithmic scales. The logistic function has three parameters: the saturation level S_{max}, the parameter that situates the change in time t_0, and the parameter that defines the speed of the change *a*. The saturation level S_{max} is usually scalable to 1.[1] The parameter that situates the process in time, t_0, or *actuation time*, is the curving point in figure (6). On a logarithmic plot, where the S curve is a straight line, it is the *intercept*. Variation in actuation time causes a horizontal shift, illustrated with a linear plot in (8).



(8)

The parameter that defines the speed of the change, or *rate* in Kroch 1989, is the steepness *a* of the S curve. On a logarithmic plot it is the slope of the straight line. In (9) we illustrate variation in *a* in a linear plot.



(9)

In the next section we dicuss these parameters in relation to the Constant Rate Hypothesis.

3.1 Kroch's Constant Rate Hypothesis

A change in the grammar does not always show up in the E-language in all contexts at the same time. Usually a change starts in a certain context, and extends its scope to other contexts step by step, until it comes to completion. A famous example is *do*-support in the history of English, which starts in negative WH questions and then proceeds to other contexts. Negative declaratives are the last context affected. Kroch (1989) suggests that changes in the E-language that are manifestations of one and the same change in the I-language share a property: the speed with which they proceed (their *rates*) is equal although their actuation times may be distinct. In terms of the logistic function in (7), it turns out that the various curves share the slope parameter *a*, while their actuation times t_0 are different. In other words,

[1] By deviding both members by S_{max}. But this is not always possible, for instance if the change from A to B involves an increase in use of the construction. If so, one does not really know when the change has completed.

the various curves in (8) may be manifestations of the same change in parameter, while the various curves in (9) cannot. This is the Constant Rate Hypothesis (CRH).[2] The CRH has turned out a major crowbar in the analysis of linguistic change. However, it has not been derived from deeper principles. Let us finish this section with a summary of the advantageous and disadvantages of the logistic model.

(10) *Advantages of Kroch's logistic model*
- Simple logistic two- or three-parameter model
- Well-known from life sciences
- Can be easily displayed and graphically manipulated
- Constant Rate Hypothesis ties E-language process to the I-language (P&P model)

There are also disadvantages of this model. The first disadvantage is the subject of this talk: the relation with transient states under S-curves cannot be made, as was first noticed by Anthony Warner: "(...) the "Constant Rate Effect" (...) (has) the difficulty that the relationship between affirmative declarative *do* and the other contexts is not clear" (Warner 2006:49). This is related to another property of the logistic model: it is a rigid three- or two-parameter model that cannot be modified in a non arbitrary way. The disadvantageous are listed under (11).

(11) *Disadvantages of Kroch's logistic model*
- Relation with the Transient State is not expressed
- Constant Rate Hypothesis remains a stipulation
- The model cannot be modified

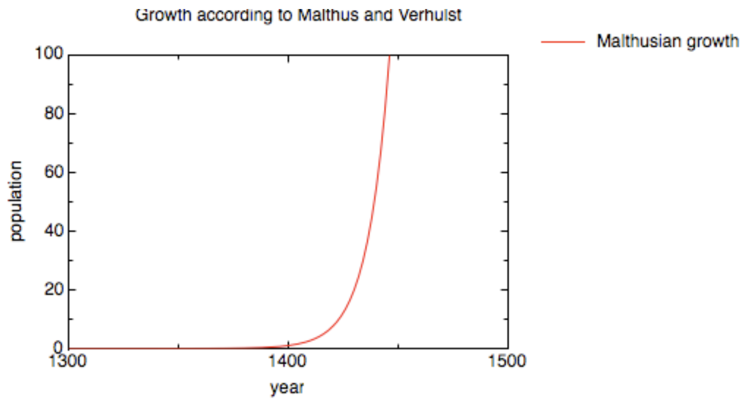
In the next section, we sketch how we can transcend these limitations without losing the advantageous features.

3.2 The underlying Differential Equation

As said, Kroch's CRH is an empirical generalization. It is not (yet) derived from underlying principles. At this point, it must be noticed that logistic curves describe the kinematics of a change in the time. It does not describe the actual process behind the change: how do the various players interact dynamically? It may be clear that a description of the outer manifestation cannot give insight in connection between these kinematics. In this section we show how a shift from functions (describing the kinematics of a change) to Differential Equations, describing the fundamental interactions, will shed light on the CRH. To illustrate the step, let us first take a simple case of Malthusian growth, as described kinematically by an exponential function under (12).

[2] The CRH is a misnomer. It should be called the Equal Rate Hypothesis. We nevertheless stick to the traditional terminology.

(12)



As Malthus has shown in 1826, the exponential growth is a consequence of a fundamental property of life: the increase of a population S is proportional to the size of the population S. The bigger the population, the more it grows. This can be expressed mathematically by a differential equation: the change rate in a quantity S is proportional to S, as given in (13a). This differential equation has a solution or rather as bundle of solutions, as given in (13b).

(13) a. $$\frac{\mathrm{d}S}{\mathrm{d}t} = aS$$ (Malthus 1826)

b. $$S(t) = e^{a.(t-t_0)}$$

This is the exponential function. For every parameter *a* in (13a), there is a set of time shifted solutions, defined by the various t_0. This model of exponential growth was modified by Verhulst in 1838. Verhulst noticed that there is a counterforce that limits the growth. The more a population grows, the less commodities there are to serve it. Verhulst, therefore, added a factor (1-S) to the equation in (13a), which reduces the growth when S becomes bigger. The new differential equation is given in (14a). It solution is an S-curve, the logistic function. It has, once again, not one solution, but a bundle of time-shifted solutions, given in (14b).

(14). a. $$\frac{\mathrm{d}S}{\mathrm{d}t} = aS(1-S)$$ (Verhulst 1838)

b. $$S(t) = \frac{1}{(1+e^{-\frac{(t-t_0)}{a}})}$$

Verhulst's strategy may be clear: DEs have *interpretations,* which are usually rather straightforward. This makes them modifiable in a non-arbitrary way. One argues against or in favor of each factor or term in the DE. And subsequently one finds an analytic solution or a solution by computer simulation. Essential is, therefore, the interpretation of the differential equation.

Let us now return to the language change of A → B, which is described by an S-curve (14b). The related DE in (14a) has at least two interpretations that are relevant for the present discussion. There is the interpretation by Malthus: apart from the factor that stimulate a change, say the wish for innovation, there is a counterforce, say, social acceptability or whatever. These two factors interact in a certain way. Adding (1-S) leads to (14a) but it can in principle be modified by another type of factor. The other interpretation is a more statistical one. If S is the number of people that have shifted from A to the innovative B, 1-S is the number of people that has not undergone the change. We therefore can give the following interpretation to the underlying DE in (14a). Let us assume that speakers influence each other when they meet. If a B-speaker meets an A speaker, A may change to B. The increase in the number of speakers B is proportional to the probability that a speaker of B meets a person that is still an A-speaker. If we assume the extremely simple model where all speakers meet each other with equal probability, the DE in (14a) holds.

Notice that the DE in (14a) is the underlying equation that has the logistic function as its solution. Notice further that this solution is not unique but has a bundle of time-shifted solutions. The constant *a* in (14b) is fixed by *a* in (14a), but t_0 in (14b) is completely free. So, if the equation in (14a) describes a language change, we predict many solutions that instantiate this change. These may have different actuation times, t_0, but must have equal slopes or rates, *a*. This derives Kroch's Constant Rate Hypothesis. We also now know, when we may expect the CRH to hold: only in the case that all (relevant) speakers meet each other by equal chance, i.e. without geographical effects and without diffusion, is (14a) the applicable DE and does the CHR hold. In the next section we apply this promising method to the transient state. We may summarize the conditions of the logistic model and the CRH as in (15).

(15) <u>Conditions of the logistic model</u>
The DE underlying the logistic model (Verhulst's equantion) only holds iff
- there are just two variants in competition: A and not-A
- there is only a first-order interaction, i.e. no inherent decline/ increase (e.g. increase of literacy might favor a certain construction)
- everyone meets with everyone with equal chance (no geographical effects, diffusion, etc.)

In the next section we modify the underlying DE of the logistic model with the presence of a transient state.

4. The 3-state model

In the previous section we reported a well-established method from physics and the life science to modify processes in a non-arbitrary way. One adds a well-interpreted term or factor to the ruling differential equation and solves the new equation analytically or by computer simulation. In this section we apply it to failed changes.

4.1 The failed change model (Postma 2010).

A failed change is defined as a combination of a monotone increasing logist and a subsequent monotone decreasing logist with the same slope. A change is "inherently failing" if the actuation time of the increasing and decreasing logist coincide. If we denote the increasing logist with A, then the decreasing logist with equal actuation time and opposite slope is (1-A). The failed change is therefore A(1-A), and by using the DE in (14a), repeated here under 16), we conclude that the failed change F=A(1-A) is proportional to the first derivative of the successful change.

(16) $$\frac{\mathrm{d}A}{\mathrm{d}t} = aA(1-A)$$

It has turned out that this model give quite precise predictions of the transient state that occurs under various successful changes. Not only the example given above of the rise of reflexive in

Middle Drentish is nicely described (cf. 6), also the failed change of DO-support in positive affirmative contexts comes out rather precise. The defect of this model is that the failed and successful change are not described by one set of (coupled) DEs, but by one DE which is interpreted in two different ways. In the next section we will correct that by developing the transient state model.

4.2 Preparatory step: transition diagrams

Verhulst's equation describes the transition of A → B by describing the rise of B or the decline of A together with a conservation law of speakers: A + B=constant. In the model no persons die, dissolve in the air, or become non-speakers. This can also be expressed by two coupled differential equations as in (17), which is fully equivalent to (16).

(17)
$$\frac{\mathrm{d}A}{\mathrm{d}t} = -a.A.B$$
$$\frac{\mathrm{d}B}{\mathrm{d}t} = a.A.B$$

The equivalence can be inspected by taking the sum of both members, resulting in d(A+B)/dt =0, which states that the change in A+B is 0, i.e. A+B is constant. By substituting B=1-A in the first equation of (17) we derive (16). Similarly we can go from (16) by substitution of 1-A=B in the first equations in (17).

A two-state change A → B can, hence, be expressed by two *coupled* DE. The equations in (17) have an interesting interpretation. The transition away from A and towards B comes about by the interaction of A and B, i.e. the chance that A meets B, multiplied by the chance *a* that this encounter leads to a change from A to B, i.e. the transition coefficient *a*. In this case the relation between the two equations in (17) is trivial: the amount by which A diminishes is equal to the amount by which B increases. We can visualize it in the *transition diagram* in (18).

(18) *Transition diagram* or *Transition diagram*

$$A \xrightarrow[\alpha A.B]{} B \qquad \text{or} \qquad A \xrightarrow[\alpha]{} B$$

In the next section, we apply this method to the three-state model.

4.3 Modeling transient states

Let us assume a simple stable model without overall increase of decrease in speakers and where children simply replace their parents in society numerically and linguistically. So, if the child enters society, we consider the parent to be removed. This is a rough approximation which excludes interactions between generations in society. By using the techniques developed in the previous section, we can model a three-state change A → T → B by assuming the transition diagram in (19).

(19) *Transition diagram*

$$A \xrightarrow[\beta A.T]{} T \xrightarrow[\alpha T.B]{} B$$

This leads to the coupled differential equations in (20).

(20)
$$\frac{\mathrm{d}A}{\mathrm{d}t} = -\beta A.T$$
$$\frac{\mathrm{d}T}{\mathrm{d}t} = +\beta A.T - \alpha T.T - \alpha T.B$$
$$\frac{\mathrm{d}B}{\mathrm{d}t} = \qquad +\alpha T.T + \alpha T.B$$

Let us interpret the terms. The term A.T is the chance of an encounter of an A-speaker and an innovator T (say the *sick* speaker in our example), which has changed the language with a peripheral rule. If these two speakers meet, speaker A turns in an innovator himself with chance β. We may call this an L2-interaction. The term T.B turns an innovator into a B-speaker. This involves a parameter reset, which is only possible upon L1 acquisition, i.e when the child replaces the adult. One should keep in mind that the innovation is not really grammatical, as was assumed in the Weerman model, and upon this L1-interaction, the innovator cannot completely resist turning into the more grammatical B (*sich* in the case of the example). Finally, there is the term T.T which represents the encounter of an innovator with another innovator. To understand what this means, it is important to notice that we assume that nothing happens whenever A meets A, or whenever B meets B: no terms A.A or B.B are included. These are grammatical utterances in a homogeneous speech interaction, which do not cause a change. This is different, however, in the case of a T-T encounter. It is generally assumed in innovation/rumor theory that innovators stop using the innovation as soon they meet another innovator (Piqueira 2010). -- As if a woman in a Viktor & Rolf dress meets another woman in a similar dress. She will never wear it again. -- In the case of language change we must split this term into an adult-adult interaction and an adult-child interaction (or rather child-adult replacement in our model). The adult-adult interaction does not contribute a change as adults cannot reset parameters, which is needed to switch to B. The adult-child interaction, or rather, when the child replaces the parent in society, results in B. A similar reasoning holds for the interaction between T and B. Switching to B needs a child. So, the transition coefficient β involves L2-interaction, and the transition coefficient α involves L1. So we fundamentally have built in the transient intermediate state with its properties into the model. This model has been described and computer-simulated in Piqueira (2010). It is an implementation of the famous SIR-model[3] in biology modeling, the spreading of diseases (Kermack & McKendrick 1927).[4]

The coupled differential equations in (20) are not solvable by analytic methods. We have solved them using numerical methods.[5] The solution is displayed in the figure in (21).[6]

---

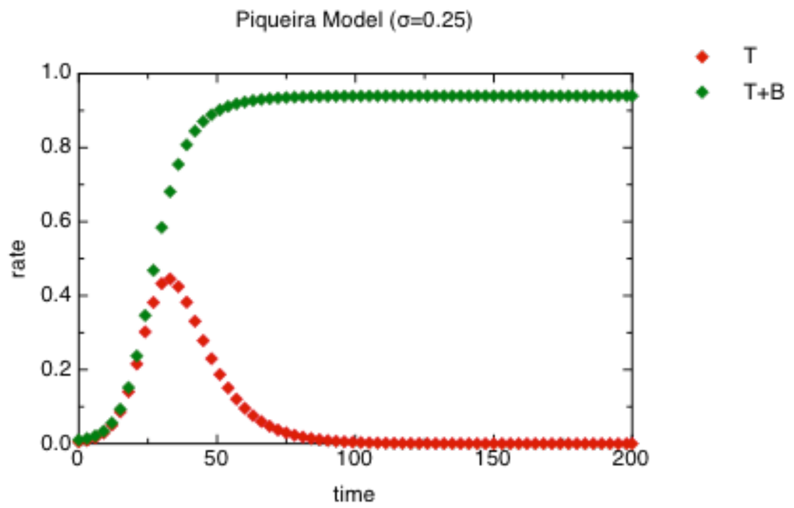
[3] Susceptible-Infected-Removed in the epidemological model.

[4] Piqueira's model is an application of the biological model to the spreading of information: the Daley-Kendall model. The three groups are called Stiflers, Spreaders, and Ignorants. The differential equations in Piqueira's 2.1 and ours (cf. 20) are identical, modulo the error in Piqueira's 2.1).

[5] The Runge-Kutta method. We used the program SolveDiffEq 3.6 for MacOS by Bob Delaney's Science Software 2013.

[6] SolveDiffEq does not have a graphical component. We used ProFit 6.2.11 for MacOS.

(21)



We see a rather similar picture to the figure in (6), which displays the failed change of *sick* in relation to the successful change of the s-reflexive in the failed change model, where the failed change is first derivative (or slope function) of the successful change. In the new model, however, this relation is an approximation. Notice that (21) is not a fit to the historical data, which we have not performed yet. It would mean an in tandem solving of differential equations and a fitting process. We hope to carry that out in further research. What the present discussion tells us, is that we succeeded in making a consistent model of a realistic linguistic process that includes the successful and the failed change (transient state), where the transient state is the initiator of the successful change.

A final comment on (21) is in place. It must be noticed that the T+B line does not reach the full saturation level of 100%. This means that a small number of A-speakers remain after the T-innovation has swept through the community. The reason is that we did not include a term of direct interaction between A and B, say $\gamma A.B$. If we add such a term with negative sign to (20a) and with positive sign to (20c), full completion of the change is guaranteed. However, this simply superposes the new model with the terms of the old model (17). These terms are inherently asymmetric. Alternatively, one might add a symmetric change like A.B.B → B + A.A.B→A to the system. This three-individual interaction would add an effect of accommodation to the majority. In terms of the equations in (20) it would add a term AB(A-B) to (20a) and AB(B-A) to (20c). However, as these changes can only happen within the family (from A to B is a parameter switch), we must assume core families with more than two adults.

5. Conclusions

It is possible to generalize the logistic model of language change from a kinematic modeling of the change to a dynamic differential equation that rules the change: the diferential equation captures the respective interactions and functions as a process *law*. In the simplest case, this law takes the form of Verhulst's equation. The solution of this equation is a logistic function, or rather a bundle of the logistic functions, which are time-shifted with respect to each other. This derives the Constant Rate Hypothesis. By modifying Verhulst's equation with an intermediate state, we made a unified model (coupled differential equations) that models both the successful change and the failed change within one model. The failed change is the first derivative of the successful change, by approximation only.

6. References

Altmann, G. *et al.* (1983). A Law of Change in Language. In: B. Brainard (ed.). *Historical Linguistics*. Bochum, West Germany: Studienverlag Brockmeyer. 104-115.

Andersen, Henning (1973). Abductive and Deductive Change. *Language* 44. 765-793.

Andersen, Stephen, and David Lightfoot. (2002). The language organ: Linguistics as cognitive physiology. Cambridge, U.K.; New York: Cambridge University Press.

Bailey, Charles-James. (1973). *Variation and Linguistic Theory.* Washington: Center for Applied Linguistics.

Bybee & Slobin (1981). Why small children cannot change language on their own – evidence from the past tense. In: Ahlqvist, Anders (ed.), Papers from the Fifth International Conference on Historical Linguistics, Galway, April 6–10 1981, 29 ff.

Chomsky, Noam. (1965). Aspects of the Theory of Syntax. Cambridge: MIT Press.

Chomsky, Noam. (1986). *Knowledge of Language: Its Nature, Origin, and Use.* Greenwood.

Kermack, W. O. and A. G. McKendrick. (1927). "Contributions of mathematical theory to epidemics," Proceedings of the Royal Society Series A. 115. 700–721.

Kroch, Anthony (1989). Reflexes of Grammar in Patterns of Language Change *Language Variation and Change,*1989, 1:199-244.

Lightfoot, David W. (1999). *The development of language: acquisition, change, and evolution.* Oxford: Blackwell.

Pereira, José Roberto (2010). Rumor Propagation Model: An Equilibrium Study. *Mathematical Problems in Engineering,* Article ID 631357.

Postma, Gertjan (2004). Language contact and linguistic complexity - the rise of the reflexive pronoun 'zich' in a 15th century Netherlands' border dialect. In: Stephen Anderson & Dianne Jonas (eds.) *Syntactic Variation and Change, Proceedings of DIGS-8 at Yale*. Cambridge University Press.

Postma, Gertjan (2010). The impact of Failed Changes. In C. Lucas, S. Watts, A. Breitbarth, & D. Willis (Eds), *Continuity and Change in Grammar.* Benjamins. 269-302.

Rogers, E.M. (1962). *Diffusion of Innovations*. Free Press. New York.

Sankoff, Gillian & Helene Blondeau (2007). Language change across the lifespan: /r/ in Montreal French. ICHL Montreal.

Anthony Warner (2006). Variation and the Interpretation of Change in Periphrastic do. In. Ans van Kemenade & Bettelou Los (eds). *The Handbook of the History of English*. Blackwell. p.45-91.

Weerman, Fred (1993). 'The Diachronic Consequences of First and Second Language Acquisition: the Change from OV to VO', *Linguistics* 31, p.903-931.

Wexler, Kenneth (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106: 23-79.

Van der Wurff, W. (1990). *Diffusion and Reanalysis in Syntax*. Doctoral dissertation, Leiden University.

Weinreich, Uriel, W. Labov and M. I. Herzog. (1968). Empirical Foundations for a Theory ofLanguage Change. In: *Directions for Historical Linguistics: A Symposium.* Edited by W. P. Lehmann. 95-195. Austin: University of Texas Press.