



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

The impact of failed changes

Postma, G.J.

published in

Continuity and Change in Grammar
2010

document version

Publisher's PDF, also known as Version of record

document license

CC BY

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Postma, G. J. (2010). The impact of failed changes. In C. Lucas, S. Watts, A. Breitbarth, & D. Willis (Eds.), *Continuity and Change in Grammar* (pp. 269-302). Benjamins.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

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

E-mail address:

pure@knaw.nl

Linguistik Aktuell/Linguistics Today (LA)

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

General Editors

Werner Abraham
University of Vienna /
Rijksuniversiteit Groningen

Elly van Gelderen
Arizona State University

Advisory Editorial Board

Josef Bayer
University of Konstanz

Christer Platzack
University of Lund

Cedric Boeckx
ICREA/Universitat Autònoma de Barcelona

Ian Roberts
University of Cambridge

Guglielmo Cinque
University of Venice

Lisa deMena Travis
McGill University

Liliane Haegeman
University of Ghent

Sten Vikner
University of Aarhus

Hubert Haider
University of Salzburg

C. Jan-Wouter Zwart
University of Groningen

Terje Lohndal
University of Maryland

Volume 159

Continuity and Change in Grammar
Edited by Anne Breitbarth, Christopher Lucas, Sheila Watts and David Willis

Continuity and Change in Grammar

Edited by

Anne Breitbarth
Ghent University

Christopher Lucas
Sheila Watts
David Willis
University of Cambridge

John Benjamins Publishing Company
Amsterdam / Philadelphia

™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Continuity and change in grammar / edited by Anne Breitbarth...[et al.].

p. cm. (Linguistik Aktuell/Linguistics Today, ISSN 0166-0829 ; v. 159)

Includes bibliographical references and index.

1. Language and languages--Variation. 2. Grammar, Comparative and general. 3. Sociolinguistics. I. Breitbarth, Anne, 1976-

P120.V37C667 2010

417.2--dc22

2010014516

ISBN 978 90 272 5542 6 (Hb ; alk. paper)

ISBN 978 90 272 8807 3 (Eb)

© 2010 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

List of contributors	VII
Introduction: Continuity and change in grammar <i>Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis</i>	1
PART I. Continuity	
What changed where? A plea for the re-evaluation of dialectal evidence <i>Katrin Axel & Helmut Weiß</i>	13
Impossible changes and impossible borrowings: The Final-over-Final Constraint <i>Theresa Biberauer, Michelle Sheehan & Glenda Newton</i>	35
Continuity is change: The long tail of Jespersen's cycle in Flemish <i>Anne Breitbarth & Liliane Haegeman</i>	61
Using the Matrix Language Frame model to measure the extent of word-order convergence in Welsh-English bilingual speech <i>Peredur Davies & Margaret Deuchar</i>	77
On language contact as an inhibitor of language change: The Spanish of Catalan bilinguals in Majorca <i>Andrés Enrique-Arias</i>	97
Towards notions of comparative continuity in English and French <i>Remus Gergel</i>	119
Variation, continuity and contact in Middle Norwegian and Middle Low German <i>John D. Sundquist</i>	145
PART II. Change	
Directionality in word-order change in Austronesian languages <i>Edith Aldridge</i>	169
Negative co-ordination in the history of English <i>Richard Ingham</i>	181

The impact of failed changes

Gertjan Postma

Meertens Institute, Academy of Sciences*

We study linguistic changes that failed, and their relation to successful changes. The inclusion of failed changes into Kroch's logistic model of linguistic change is possible and, in fact, necessary. The logistic functions are solutions of a differential equation that also describes how failed changes are temporal-dynamically related to successful changes: the failed change is the first derivative of the successful change. We interpret this relation in Weerman's Peripheral Rules Model: failed changes are L2 innovations supported by peripheral rules rather than by core grammar, which are subsequently accommodated and modified by L1 adopters and turned into a successful change. Two case studies support the claims: the rise of the reflexive in Dutch and the rise of *do*-support in English.

1. Introduction

Language changes. This fact of everyday life is a problem for generative linguistic theories, which, in the structuralist tradition, take language to be a *system*, composed of elements whose only status in reality lies in their relations with other elements. As such, change takes place as a replacement of one system by another, and is caused by factors *outside* the system itself. It can neither be stated nor explained within the linguistic system, that is within I-language. To the extent that the linguistic system is stable, all changes and all small deviations from the optimal realization return to the stable state, and any change would necessarily *fail*. Consequently, failures are to be expected, successes are the exception. Probably as a consequence of this reasoning, virtually no attention has been paid to the dynamics of failed changes and all attention has been focused on the dynamics of successful changes. As structuralist linguistics does not provide a theory or model for a successful change, change is considered to be the replacement of one system by the other, be this a *jump* in the individual or group of

*I am grateful to Charlotte Galves, Tony Kroch, Marc van Oostendorp, Ans van Kemenade and the audiences of the Workshop on Language Change, Amsterdam, 1 Nov 2007, the Theelezing, Meertens Institute Amsterdam, 20 Dec 2007, the Workshop on Continuity and Change, Cambridge, 18–20 March 2008, and the Diachronic Generative Syntax Workshop, Campinas, 22–24 July 2009 for discussions and comments.

individuals as an adult innovation (L2 model of linguistic change), or a *jump* between individuals as incomplete or imperfect intergenerational transmission of language (L1 model). Therefore, theories that describe the dynamics of the change necessarily involve the realization and behaviour of the E-language.

In this study we explore the properties of failed changes in linguistics and propose a model for them. It is our aim to study diffusion that goes hand in hand with a change of the system, and in order not to complicate things beyond reason, we study the simplest case of diffusion in a community as a whole, ignoring spatial or sociological effects.¹ We do so using a case study from Middle Dutch syntax and compare it with the scarce data on failures in the literature. It turns out that failed changes are related to successful changes in a systematic way. This relation can be expressed algebraically. Failed changes are the first derivative of the successful change. We interpret failed changes as what we propose to call ‘off-grammatical innovations’, that is, ‘tricks’, peripheral rules or stylistic rules (cf. Weerman 1993; Hinterhölzl 2004) made by adults or L2 learners.² The successful change is a response to the off-grammatical input of their predecessors and establishes a newly accommodated I-language.

First we review the existing models of successful change as proposed in the linguistic literature, especially Kroch’s (1989) logistic implementation and the Constant Rate Hypothesis (CRH) in Section 2. Then we show the shortcomings of this model, especially its non-application to failed changes. A study of the phenomena that have led Kroch to his logistic model leads us to the proposal that failed and successful changes are related, contradicting the CRH. Related changes do not have (seemingly) equal slopes. In order to confirm the relation, we construct a case where a failed and a successful change are clearly connected. This is our case study of the replacement of the reflexive use of the pronoun *hem* ‘him’ in a Middle Dutch border dialect by a lexical reflexive: *sick*, borrowed from Low German (the failed change), and *sich*, borrowed from High German (which finally succeeds) (Section 3). In Section 4 we then restore Kroch’s CRH by slightly modifying the model so that it includes failed changes. In Section 4.1 we will model the relation between failed and successful changes, derive the relation (Section 4.2) from other principles, and explain it in pre-theoretical terms (Section 4.4). As a nice consequence, we do not need to go beyond the logistic model to express the relation. In fact, *only* the logistic model is able to capture the relation between failure and success in a conceptually elegant way. The CRH, which links the

1. For diffusion over social networks, see Van der Wurff (1990).

2. We take the maturation period as the only watershed in acquisition and capture, therefore, all types of acquisition of linguistic features after the maturation period as L2, be it adult innovations in their proper mother tongue or true second language acquisition. Weerman (1993) assumes these to be distinct, which is meant only sociologically not linguistically (personal communication).

dynamics of the change to switches in the Principles and Parameter Model will be critically evaluated and replaced by a stronger version. In Section 5 we evaluate the results thus far. We come to the conclusion that the model can best be interpreted in terms of an L2-theory of linguistic change as sketched out in Weerman (1993). In Section 6 we reanalyse the data in Ellegård (1953) and Kroch (1989) in terms of the failed change model. The chapter closes with a pre-theoretical reflection on the model, its results and its consequences.

2. Successful changes

Gradual linguistic changes that can be observed in a non-homogeneous language community were studied by Weinreich, Labov & Herzog (1968). Not contradicting the Chomskyan idealization of linguistic competence as displayed in a homogeneous community, but rather complementing it, Weinreich, Labov & Herzog study the gradual shift in a non-homogeneous environment from a linguistic configuration producing A to an alternative configuration producing B through all intermediate stages. Bailey (1973) proposes a quantificational model in which change starts slowly, picks up speed in the middle and then decelerates as it proceeds to completion (see the review of Naro 1984). This describes a typical S-shaped curve as often displayed in empirical studies (Piotrovskaja & Piotrovskij 1974) and illustrated in Figure 1.

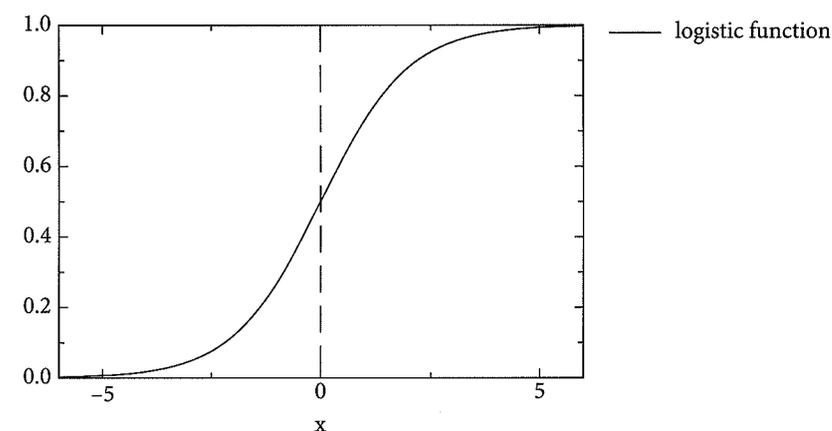


Figure 1. The logistic function as an example of an S-shaped curve

There are various mathematical functions to model such an S-curve, e.g. the Gauss cumulative function (Aldrich & Nelson 1984, Piotrovskaja & Piotrovskij 1974), the Lorentz cumulative function, and the logistic function. Altmann et al. (1983) propose a model in which the change starts out with an increase which is proportional

to the fraction of new forms, multiplied by the fraction of old forms. This establishes Verhulst's differential equation:

$$(1) \quad dS/dt = S \cdot (1-S) \quad (\text{Verhulst differential equation})$$

$$(2) \quad S(t) = \frac{1}{(1+e^{-t})} \quad (\text{logistic function})$$

The left-hand side of the equation in (1) represents the growth in a population S per time unit (dS represents the infinitesimal change in S , dt the infinitesimal change in t). The equation claims that the growth is proportional to two factors, S and $(1-S)$, to be discussed below. The differential equation in (1) has the logistic function of (2) as its solution. This means that populations that satisfy equation (1) grow with a rate represented by the logistic function.

This differential equation was first proposed by Verhulst in 1838 as an alternative to the differential equation of Malthusian growth, $dS/dt = S$, in which the growth in S is proportional to S . The latter captures the concept of unrestricted procreation. It gives an exponential growth curve ($S = e^t$). The Verhulst equation in (3), on the other hand, has an additional saturation factor $(1-S)$, which can be taken as a counterforce to unlimited growth. Instances of this counterforce are the decrease in food or space or some other factor with limited availability. In diffusion processes of linguistic innovations, the two terms may receive a probabilistic interpretation: the change for a switch in usage from A to B is proportional to the number of speakers who use B (S), multiplied by the chance of them meeting speakers that do not use B yet $(1-S)$. This latter factor, of course, decreases and forms the saturating factor. In this way, Altmann et al. (1983) provide a specific S-curve, the logistic curve, as a model of linguistic change. This model is basically a two-parameter model. The actuation time (t_0) can be modified leading to the term $(t-t_0)$ and the steepness of the change (a) can be modified, leading to $(t-t_0)/a$.³

$$(3) \quad S(t) = \frac{1}{\left(1 + e^{-\frac{t-t_0}{a}}\right)} \quad (\text{logistic model})$$

In the rest of this chapter we use the logistic model.

2.1 Kroch's model

Kroch (1989) adopts the logistic model without much discussion. For Kroch, any two-parameter S-shaped curve suffices (such as the logistic function, the Gauss cumulative

3. The actuation parameter in a model is that parameter that defines the process's orientation on the time scale. The saturation rate is a third parameter but it is normalized to a dimensionless 1 (100%).

function, or the Lorentz cumulative function). Kroch remains agnostic as to which model fits best: "these generally differ so little from the logistic that they can provide no improvement in fit to empirical data." This is correct. Kroch adopts the logistic because of its widespread use in the life sciences and because of its graphical properties: when the occurrence rate is displayed on a semi-logarithmic graph against the time axis, straight lines are obtained.⁴ In Section 4.2 we propose theoretical reasons for exclusively adopting the logistic model derived from theoretical considerations of failed changes. The innovative part of Kroch's study is that it makes a connection between the gradual diffusion of a change in the E-language, as described by the logistic curve, and more abstract properties of the I-language such as the parameter switch in the grammatical system. This relation is established in the so-called Constant Rate Hypothesis (CRH), which states that related changes have equal slope parameters. The CRH states that an underlying change in the parameter setting may manifest itself in various constructions with different actuation times but must proceed in these constructions with equal speed. In a two-parameter model, this means that the slope parameter of the different related S-curves must be equal, but the actuation parameter may be distinct. For instance, the curves in Figure 2a with equal slope but distinct actuation times may be tied to one parameter switch (they do not need to be!), but the curves in Figure 2b, with distinct slopes and equal actuation cannot to be reduced to one parameter reset.

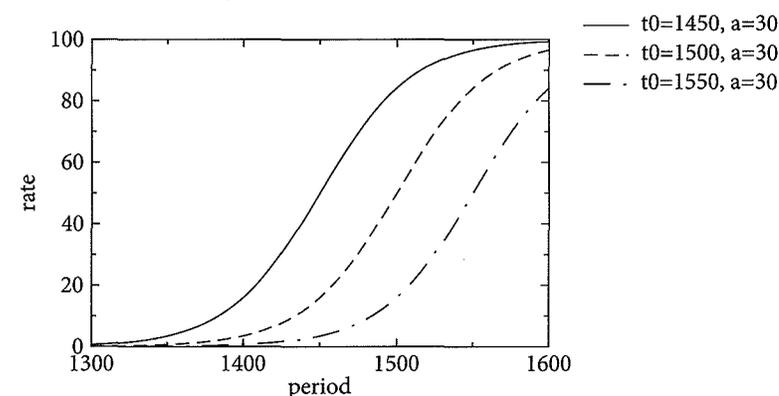


Figure 2a. Kroch's (1989) model – Logistic curves with distinct actuation times and equal slope

4. Straight lines are only obtained provided that on the logarithmic scale the Fisher-Pry transform of the occurrence rate ('odds') is used.

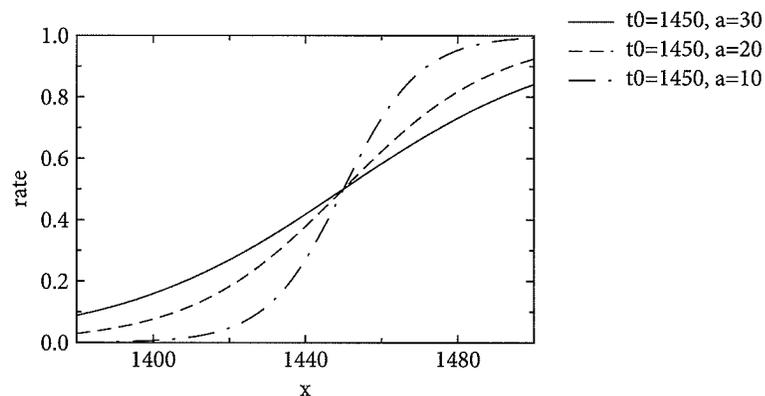


Figure 2b. Logistic curves with equal actuation times, distinct slopes

Kroch does not derive the CRH from basic principles but motivates it empirically using four case studies taken from the literature: (i) the replacement of *have* by *have got* in British English (studied in Noble 1985) in four different environments; (ii) the rise of the definite article in Portuguese possessive noun phrases (studied in Oliveira e Silva 1982) in 6 different environments; (iii) loss of verb-second word order in French in two different environments (studied in Fontaine 1985 and Priestley 1955 respectively); and (iv) the case of periphrastic *do* (“*do*-support”) in six different types of constructions, as studied by Ellegård (1953). Kroch argues that in all these cases the slopes of the various S-curves are equal, while their actuation times may be different (as in Kroch’s case studies 3 and 4) but these do not need to be (as in Kroch’s case studies 1 and 2, where both activation times and rates are equal).⁵ On the assumption that these changes are tied to one and the same parameter, Kroch derives the CRH.⁶ Its contraposition says that whenever two changes have distinct slopes, they must be tied to distinct parameter resets.

2.2 Evaluation of Kroch (1989)

Kroch’s model has been a major step in analysing linguistic change on the basis of E-language (corpora), without losing the connection to changes at the level of I-language,

5. In Kroch’s case studies 1 and 2, the stability factors (i.e. the ratio between the rates of the constructions) remain equal during the change, which implies that both their slope parameters and their actuation times must be equal.

6. Kroch uses the unfortunate term ‘constant rate hypothesis’ for the more correct ‘equal rate hypothesis’. In the logistic model any change that fits to the logistic function has a constant rate during the change. We stick to the existing term, though.

that is, tying the gradual change in the linguistic community to the discrete change at speaker level and on the level of the (individual) grammars. We take Kroch’s model as a starting point. However, there are data that are difficult to evaluate within Kroch’s logistic model, simply because of the fact that their diffusion rates do not have the shape of S-curves: some have bell-shaped curves. Nevertheless, Kroch simply treats them as having a logistic pattern. For instance, within the study of *do*-support in English, the diffusion of *do*-support in positive declarative clauses is first forced into the logistic model, two parameters are extracted, one of which, the slope, turns out to be significantly distinct. Subsequently, under application of the CRH, it is concluded that this process is parametrically disconnected from the other instances of the diffusion of *do*-support. The logistic fits are given in Figure 3 as straight lines.⁷ At first glance, positive *do*-support (the black line) isolates itself as a process separated from the other lines.

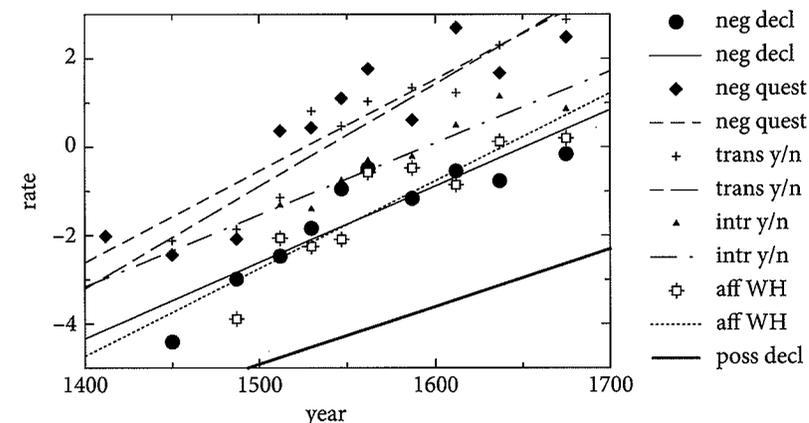


Figure 3. The rise of *do*-support in English on a logarithmic scale (data from Ellegård 1953)

Some provisos must be made here, however. In the first place, the separation of the two sets of lines is not established on the basis of rates obtained by significant fits. The black line in particular is a bad logistic fit to the underlying data. Moreover, it is unattractive to disconnect the failed process in time from the first actuation of *do*-support, as they obviously occur more or less in the same timespan, as Kroch notices himself

7. The vertical axis is logarithmic in addition to a Fisher-Pry transformation of the data. The regression calculations throughout this chapter are performed using Pro-Fit 6.1.11 for Mac. The algorithm is Levenberg-Marquardt.

(1989: 29). This can be more clearly seen from Figure 4 without a logarithmic scale, where the successful changes are approximated by S-curves and the failed change by a significant fit to a bell-curve.

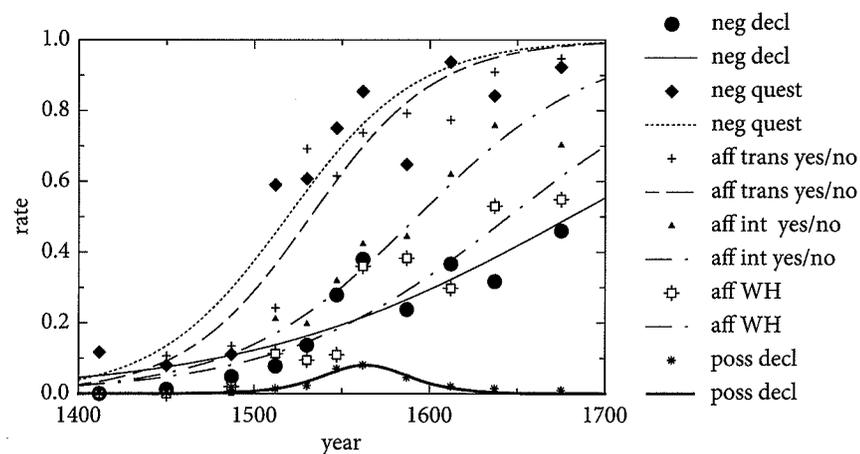


Figure 4. The rise of *do*-support in English on a linear scale (data from Ellegård 1953)

With respect to the actuation parameter, the first successful change(s) and the failed change seem to be related somehow: the peak of the bell curve coincides with the inflection point of the initial S-curves. Their activation times are equal.⁸ This cannot be expressed in Kroch's theory, as the CRH does not connect processes on the basis of an equal actuation parameter, but disconnects processes with unequal slope. If one assigns a slope to the black line as if it were an S-curve (as Kroch does, cf. his Table 9), a slope comes out that is unequal to the slope(s) of the successful lines. What can be done if no slope can be assigned, because it is not a good fit? Kroch suggests that some "more complex model should be applied" (1989: 29). An important feature of this "more complex model" should include his observation that "the rate of use of *do* in the affirmative declarative context rises along with the other contexts ... which suggests that the context is tied to the others". It is this programme that we intend to design in this study.

In the second case study used by Kroch, two processes are considered with respect to the increase of the determiner in Portuguese (cf. Oliveira e Silva 1982). Portuguese changes from the use of bare possessives as in *meu livro* to the use of possessive

8. Since the inflection point of the S-curve and the maximum of the peak are the only independently motivated orientation points on the time axis, we can only take these as the activation time.

constructions with definite determiners: *o meu livro* 'the my book'. A similar increase occurs in colonial Brazil, which does not persist.⁹ The Portuguese data on the increase of the use of the pronoun in the motherland and in colonial Brazil are compared.¹⁰ The data and the logistic fits are graphically displayed in Figure 5.¹¹

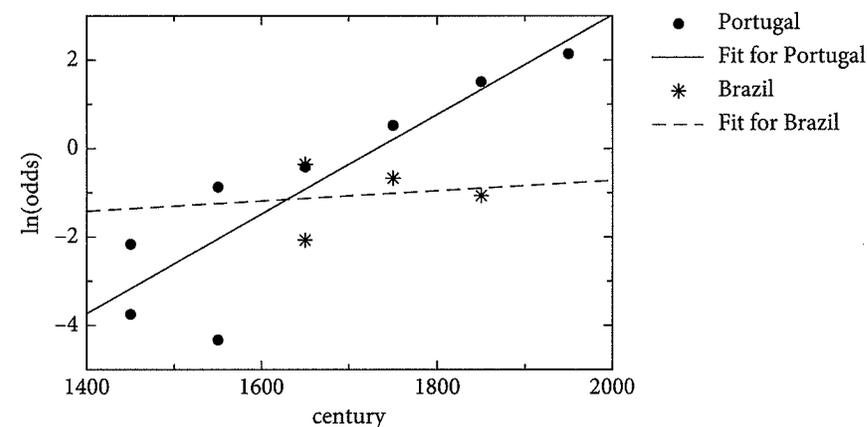


Figure 5. Logarithmic plot of the article in possessive constructions in Portuguese with S-shaped fits (data from Oliveira e Silva 1982: 314–315)

Can we conclude from these data that these changes are not connected? One should keep in mind that one of these data sets (stars/dashed line) does not have the shape of an S-curve, and is not appropriate to be fit to an S-curve. Consequently the extracted parameters do not have statistical significance. One cannot evaluate two processes on the basis of these parameters and decide on their non-relatedness. If, on the other hand, we fit the failed change to a bell curve the following picture arises.

9. The data are taken from the Rio dialect. In modern times, probably under influence of the southern dialects, a new increase of the article can be observed. We leave this later period out of consideration.

10. Oliveira e Silva includes data for the period long after the independence of Brazil. We only compare Portugal and Brazil as long as they formed a linguistic community of some kind. So we ignored the data after 1811.

11. Oliveira e Silva makes the observation that the fit is not significant.

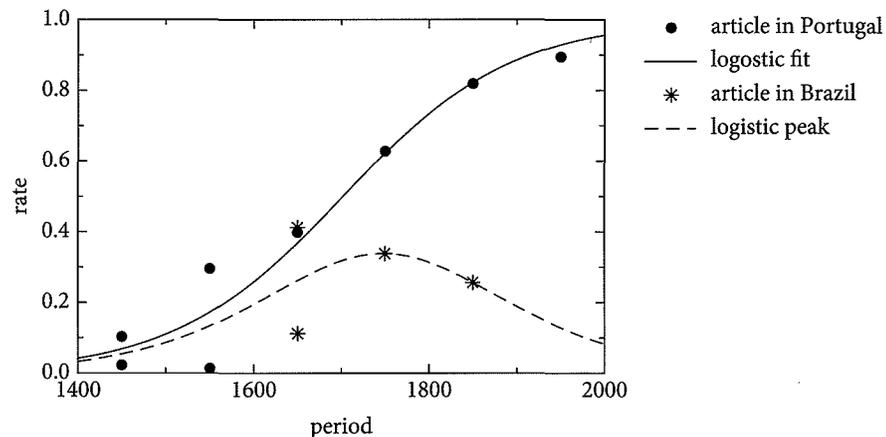


Figure 6. Linear plot of the article in possessive constructions in Portuguese in S- and bell-shaped fits (data from Oliveira e Silva 1982: 314–315)

Once again, the top of the failed change (dashed curve) seems to be related in time to the inflection point of the S-curve (drawn curve). Apparently, an attempt for change in colonial Portuguese and the success in local Portuguese are related, at least in time (i.e. in actuation).

In sum, it is suggestive that the failed change and the successful change are connected in time and possibly in shape. The top of the failed change seems, in the case of *do*-support, to be connected to the first actuation of the successful change. So the following questions arise:

- (4) – Is there a relation of some kind between a failed change and a successful change?
 – Can we construct conclusive evidence for such a relation? And if so...
 – Can we construe a model of failed and successful changes?
 – Can we connect them on the basis of (a modified form of) the CRH?

If the CRH is to be retained, we must attach some form of “slope parameter” to the bell-shaped curve of the failed change.

3. Failed changes

In the previous section we discussed two successful linguistic changes that seem to be flanked with a failed change: the maximum of the failed innovation appears to be connected in time with the inflection point of the successful innovation. In the standard logistic model enriched with the CRH, there is a problem to connect the

two innovations, since the framework does not make a connection on the basis of shared actuation times. Moreover, their slope parameters, assigned upon a fit to a logistic curve, are distinct (if assigned at all). It is plausible that the two developments are connected some way (as Kroch observes himself), although the framework is not able to express this. So, the logistic model seems to have a defect or limitation here, since application of the CRH leads to the undesirable conclusion that two processes, the successful and the failed process, are not connected. Moreover, the fit of one data set to a logistic function is not significant. Therefore, it would be good to have a conclusive case for which it can be argued on independent grounds that the failed and the successful change are connected. This can be done if the failed and the successful change show up in exactly the same context and they are in nature intertwined. This case is outlined in the next section and elaborated in the subsequent section.

3.1 Constructing a compelling case

In this section we construct a case where a failed and a successful change are connected on independent grounds. In that way, Kroch’s model of logistic diffusion + the CRH will be proved to be incomplete: if the slope parameters of the fits are distinct, we have construed a CRH-refuting case where two processes with distinct slopes are connected, in conflict with the CRH. It must be kept in mind that one of these slopes is assigned by brute force and is not significant. Subsequently, therefore, we show that in a slightly modified logistic model, a more meaningful slope parameter can be assigned to a failed change. An additional advantage will be that in this modified logistic model the CRH can be maintained.

The construction of a refuting case involves a failed and a successful change in one and the same context. For instance, where the failed and the successful innovations share a feature, call it [+s] (which is successful) but differ in a feature where one is successful and the other a failure, call it [+f].

- (5)
- | | | |
|---------------|------------|--------------|
| old situation | innovation | |
| [-s] | [+s, -f] | (successful) |
| | [+s, +f] | (failure) |

Under the assumption that the f-feature is irrelevant in the old situation and that [+s] is orthogonal to the [f] feature, it follows that [+s, +f] and [+s, -f] are connected because of [+s]. As the change from [-s] to [+s, ±f] and from [-s] to [+s, -f] are related to one parameter, one would on the basis of the CRH predict that the changes have equal rates. If we then show that the empirical data provide distinct slopes, it follows that the CRH cannot be retained in its present form, or in its present interpretation. Furthermore, the failed and successful changes have equal actuation times. So the CRH is both too weak and too strong. It is too strong as two changes with distinct slopes are related, which is excluded by the CRH. On the other hand, it is too weak: related changes (the failed and

the successful) have equal actuation times, which cannot be expressed in the theory. In the next section we work out this case. In Section 4 we develop an extension of the logistic framework that includes both successful and failed changes. This framework provides an interpretation of the CRH that makes it possible to retain it.

3.2 The refuting case: The rise of the reflexive in Middle Dutch

The case that refutes the CRH, as designed in the previous section, can be found in the rise of the reflexive in Middle Dutch, as studied in Postma (2004). Middle Dutch did not have a separate reflexive pronoun comparable to German *sich* and Latin *se/sibi*, but simply used the pronoun *hem* 'him' in reflexive contexts, just as Modern Frisian does and Old English did. Later Dutch developed a separate reflexive, which is obligatorily used in locally bound contexts. The change happened at different times for the different dialects. Earliest in the northeastern dialects, latest in Holland. In Drenthe the change happened during the fifteenth century. Using a corpus of 5000 verdicts (AD 1400–1500, 215,000 tokens) from Drenthe, a northeastern region in the Netherlands, Postma (2004) argues that, during the course of one century, new reflexive pronouns *sich* and *sick* 'himself', and their variants, collectively referred to as 's-reflexives', arise at the cost of the reflexive use of *hem*. The numerical data are in Table 1.¹²

Table 1. Reflexive contexts in 15th century Drenthe

Period	<i>sick</i>	<i>sich</i>	\emptyset	<i>hem</i>	<i>s-refl</i>	total	% <i>hem</i>	% <i>s-refl</i>	% <i>sick</i>
1403	2	1	0	33	3	36	0.895	0.104	0.08
1420	0	0	0	17	0	17	1	0.0	0.0
1435	9	6	0	18	15	33	0.568	0.431	0.25
1450	4	6	0	10	10	20	0.450	0.549	0.18
1465	11	4	0	21	15	36	0.566	0.433	0.33
1481	21	42	0	11	63	74	0.135	0.864	0.28
1497	0	24	3	12	24	36	0.325	0.675	0.0
1516	3	11	0	3	14	17	0.176	0.823	0.17

After an initial increase of *sick*, this form gradually disappears in favour of *sich* halfway through the fifteenth century. Although *sich* is successful and the *sick* form is not, it is theoretically better to say that the s-reflexive forms gradually replace the *hem*-form in reflexive contexts and that the form *sick* fails. In Figure 7 we have represented the occurrence of the form *sick* and the total rate of the s-reflexive (*sich*+*sick*) in the corpus.

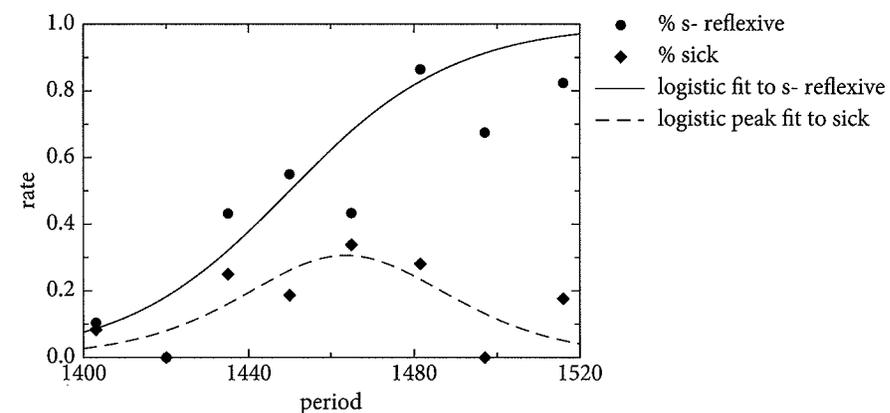


Figure 7. Rise of the s-reflexive and *sick* in 15th-century Drenthe

To connect this change to a parameter switch, it is essential to discuss the change that occurred in the I-language. In Postma (2004) I show that the change occurred as a consequence of the specialization of the pronoun *hem* going from being underspecified with respect to number ('him', 'them') to being exclusively singular ('him'). Middle Dutch *hem* had, therefore, the same underspecification as German *sich* and Latin *sibi/se*. In the course of the fifteenth century, there is a decline of the underspecification of *hem*. The requirement of referential underspecification is a formal requirement of reflexivity (Reinhart & Reuland 1993). In other words, during the fifteenth century, the pronoun *hem* loses its capability to be used in reflexive contexts. The problem was resolved by borrowing reflexive pronouns from neighbouring German dialects, which had *sick* (Low German) and *sich* (Central German), respectively. So, *hem*-to-*sick* and *hem*-to-*sich* are both related to *hem*'s loss of number neutralization. These changes are therefore related to a switch in the parameter in the I-language. These changes are related with respect to E-language as well: we see that the peak of the failed change coincides with the inflection point of the successful change. Their actuation times are related. This refutes the CRH, which is too strong in this case. Moreover, with the naked eye, it can be seen from Figure 7 that the spreading around this value, that is the speed of the switch from 0 to 1 and the speed of rise and decline, are related as well. While the steepness of the solid curve increases, the dashed curve increases, and while the solid curve starts to flatten (inflection point), the dashed curve starts to decrease (it has its maximum). At this stage, we cannot make this precise, as we lack a model for the failed change.¹³ We may, however, take these points as evidence that failed and successful

12. The division into 8 periods is obtained from the 15 period division in Postma (2004) by merging the rough data of two periods (leaving the last one untouched).

13. Anticipating the model to be developed, we give the data fitted with a logistic curve S and a logistic peak, respectively. The fitted parameters (actuation and slope) are 1452 (± 0.1) and 21 (± 0.1) for the S-curve, and 1457 (± 1.3) and 20.1 (± 0.9) for the peak, which are identical within the error.

relations are connected. The CRH cannot express this connection. It is too weak. In the next section we will propose a model for it that minimally departs from the logistic model that solves these two problems. What still has to be cleared up is the reason that *sick* failed and the reason that *sich* succeeded. That will be discussed in Section 4.5.

4. Towards a model of failed changes

In this section we model the relation between successful and failed changes. The aims are to relate successful and failed changes to each other in a systematic way. Moreover, we would like to assign *parameters* to failed changes that can be compared to Kroch's parameters of successful changes, actuation and the slope parameter.

4.1 The model

Intuitively, successful changes are S-curves and failed changes are bell-shaped curves. The question is how to model these concepts algebraically. Interestingly, S-curves and bell curves are mathematically related according to differential calculus: a bell shape is the *first derivative* of a corresponding S-curve. Alternatively, S-curves are the *integral function* of a corresponding bell curve. These relations have geometrical counterparts to which a linguistic interpretation can be given. The geometrical relations can be easily seen by inspecting Figure 8.

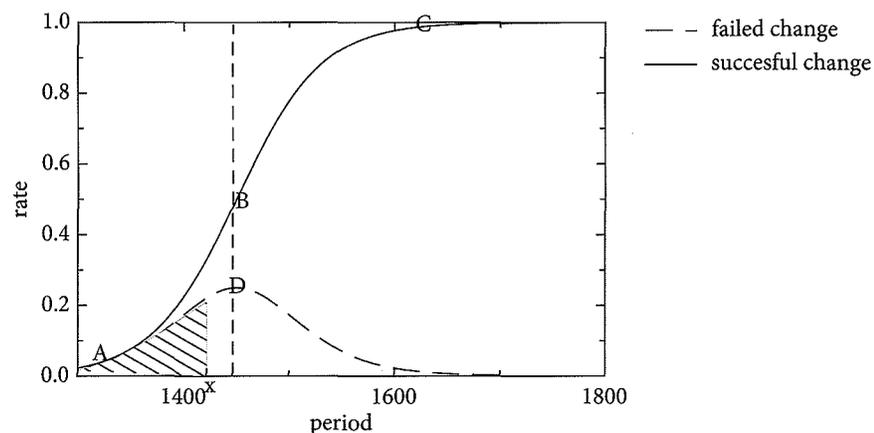


Figure 8. Arithmetic relations between failed and successful changes

Consider the subsequent *slope* values of the S-curve, rather than the values themselves. The slope starts with a zero value (at A), gradually increases until its maximal *slope* at the inflection point (at B). The slope then declines until zero again (at C). If we put the slope values into a diagram of its own, the slope function has a bell shape.

Alternatively, consider the area under the bell-curve until a certain value of the x-axis (the shaded area in Figure 8.). If the value *x* gradually increases, the area under the bell-curve that is cut off increases as well. The area grows most rapidly at the point where the bell-curve has its top (at D). Subsequently, the area continues growing but at a lower and lower rate. The S-curve is the cumulative function of the bell curve. This is true for all S-curves and corresponding bell curves.

If we now apply this to successful changes and failed changes we get the hypotheses in (6), in mathematical form in (7).

- (6) a. The failed innovation is proportional to the slope function of the successful innovation
 b. The successful innovation is proportional to the cumulative failed innovation
- (7) Let *S* be an occurrence rate in a successful change and *F* an occurrence rate in a corresponding failed change, then
- a. $dS(t)/dt = F(t)$
 b. $S(t) = \int_{-\infty}^t F(\tau)d\tau$

Although (7a) and (7b) are mathematically equivalent, they receive different linguistic interpretations. These interpretations are in fact reversed in cause and effect. We may interpret (7a) as expressing that the successful change *S* is basic and the failed change *F* a side effect: when the successful change has its highest impetus (at B), it causes resonance outside its proper domain, that is, in contexts that are less susceptible or appropriate for the change. The failed change is a spurious side effect. The formula in (7b), on the other hand, has a reversed interpretation as to cause and effect, which is slightly more difficult to grasp: it may be interpreted as a (failing) innovation proceeding autonomously for a certain period of time (because of fashion, language contact, or whatever reason). The successful change, then, is proportional to the cumulative usage of this (failing) innovation: a language community is susceptible to the number of innovative occurrences that they have heard and start to use systematically. We work out this interpretation in Section 4.4.

Thus far, the theory has had nothing to say on what type of S-curve is used in the model. The most widely used S-curves are the logistic function, the Gauss cumulative function and the Lorentz cumulative function. They are related to their respective bell curves: the logistic peak, the Gauss peak, or the Lorentz function according to the mathematical relations in (7). In the next section, we see that the logistic model is the only model that is compatible with pre-theoretical intuitions of what a failed change is.

4.2 Deriving the model: Accidental and inherent failures

In this section we derive the algebraic relations from basic intuitions of what a failed change is. What do we mean by a failed change? Now, nothing is eternal in this world, so even the most successful innovation will decline after a certain period and will be replaced by something else. We do not want to call the type of curve rendered in

Figure 9a a failure, simply because the innovation disappears after some time. Changes that, after a certain period of reasonable success, turn out to be unsuccessful and decline, simply do exist. They should not be considered failures. Such an accidentally failing change, AF, we can model as $AF = S_a \cdot (1-S_b)$, with a and b representing distinct actuation times of the rise and the decline. On the other hand we have changes whose decline is intimately related to their rise, that is, whose decline is inherent and unavoidable. They are *inherent failures* or *nonstarters*. We rather mean the latter type when we speak in this chapter of 'failed changes.' For accidental failures, the actuation of the decline is (much) later than the actuation of its rise, as displayed in Figure 9a.

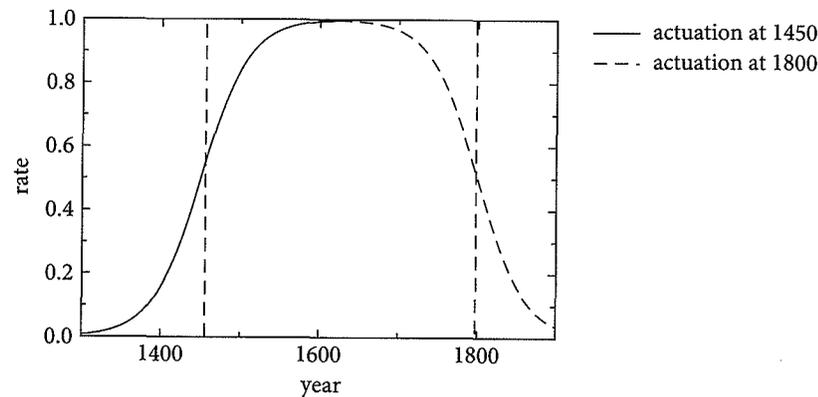


Figure 9a. Representation of an accidentally failed change: composed of two S-curves with remote actuation times

The actuation times are indicated with dashed lines. The basic intuition now is that the more the actuation times of the rise and fall come closer to each other, that is, the more the dashed lines come closer, the more 'failed' the change is, as in Figure 9b.

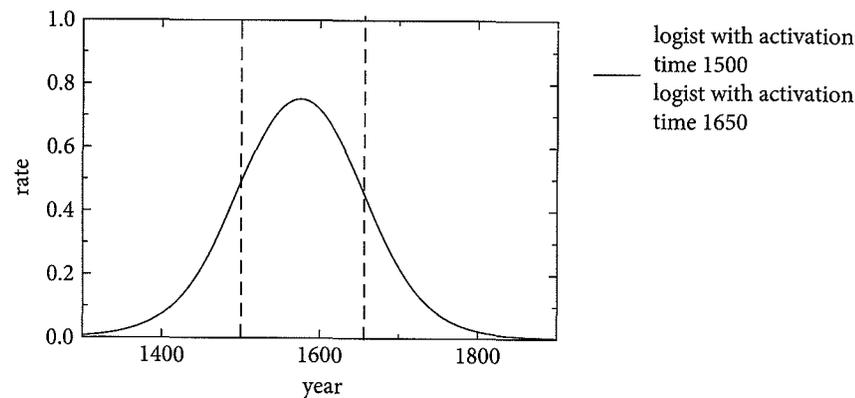


Figure 9b. Representation of an accidentally failed change: composed of two S-curves with close actuation times

If the two actuation times coincide, that is, the innovation starts declining from the moment of its rise, the change is an *inherent failure*. From this perspective, the logistic model allows a change that includes from its very beginning the ingredient of failure, but nevertheless obtains some non-zero value. Put differently, although the failed change does not exist in the I-language, it nevertheless shows up in the E-language, albeit ephemerally. This situation is displayed in Figure 9c, where the two actuation points coincide.

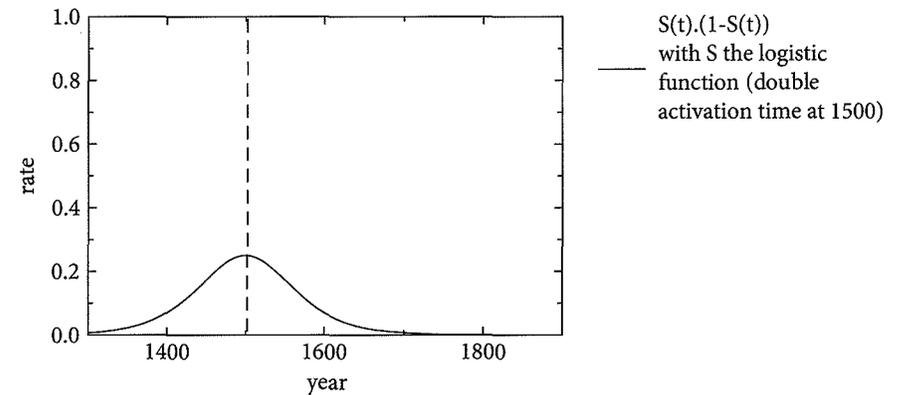


Figure 9c. Representation of an inherently failed change: composed of two S-curves with equal actuation times

Let us, therefore, define a failed change as a change that is composed of two factors, a rising S-curve multiplied by a falling S-curve, whose decline has an activation parameter equal to its rise. We make a similar assumption concerning the slope parameter, since this assumption constitutes the simplest model with the fewest parameters possible. We come back to this choice in the next section. So let us assume that a failed change is a change whose actuation and slope parameters are equal to those of its decline, that is, $F_a = S_a \cdot (1-S_a)$ with two S-curves with equal parameter sets. We call S_a the *constituting function* of the failed change F_a .

Now that we have defined a failed change, we can see what predictions it gives in (and only in) the logistic model. In the logistic model, the S-curve is a solution of Verhulst's differential equation given in (1), repeated in (8).

$$(8) \quad dS/dt = S \cdot (1 - S) \quad (\text{Verhulst 1838})$$

Now, the right-hand member of (8) is identical to a failed change, as we just have defined it: $F = S(1-S)$. Consequently the failed change is proportional to the first derivative of a successful change that constitutes it. The hypothesis is now that the constituting S of the failed change *de facto* is realized in the E-language as the flanking successful change. In other words, Verhulst's equation in (8) establishes an algebraic connection of the failed and the successful change. The successful change is the solution of the equation (S, the logistic function). The failed change is shown by the equation. Both the left and the right

member show the algebraic expression of the failed change: it is the first derivative of the successful change, as well as a superposition of two S-curves, one going up (S) and one going down (1-S). In (9) we give the algebraic shape of this failed change. It is the first derivative of the logistic curve. It is also known as the Hubbert curve.¹⁴

$$(9) \quad F(t) = \frac{e^{-\frac{(t-t_0)}{a}}}{\left(1 + e^{-\frac{(t-t_0)}{a}}\right)^2}$$

In Section 4.4 we give a conceptual justification for the model. For the time being, we are only concerned with the empirical justification. Let us, therefore, provide best fits to the failed and successful trajectories of the various reflexive pronouns in Drenthe, and use the logistic function as given in (2) and the logistic peak as given in (9). We then we arrive at the following values for the parameters, $t_0 = a[1]$ and $a = a[2]$.

Table 2. Parameter fits by the logistic function and the logistic peak to the rates of the s-reflexive and *sick*, respectively

s-reflexive (Chi squared = 0.1369)		<i>sick</i> (Chi squared = 5.8851e-2)	
Parameters	Standard deviations	Parameters	Standard deviations
a[1] = 1458.25	$\Delta a[1] = 7.9$	a[1] = 1463.6411	$\Delta a[1] = 7.9816$
a[2] = 28.1549	$\Delta a[2] = 8.4$	a[2] = 16.9110	$\Delta a[2] = 5.9411$
		a[3] = 1.2251	$\Delta a[3] = 0.3$

The two fits are made in separate fitting procedures. As we see, the a[1] parameters which represent t_0 , that is, the curving point, and the peak value of the failed change (1458 and 1463 respectively), are equal within the error. Similarly, the a[2] parameters, which represent a , that is, the steepness of the successful change, and the failed change, are also equal within the error (28±8 and 17±6 respectively). The latter identity confirms the CRH applied to failed changes. The former is the additional identity of the actuation times of the failed and the successful changes, which identity is not predicted by the CRH but has its natural place in our model.

4.3 The symmetry of the failed change

In the previous section, we modelled the failed change as the superposition of two successful changes with equal actuation parameters (by virtue of being an *inherent* failure) and equal slope parameters (for the sake of simplicity), algebraically: $F_a = S_a \cdot (1-S_a)$. This assumption causes the peak curve of the failed change to be *symmetric*. A similar

assumption of (point) symmetry is commonly made about the shape of S-curves (the initial rise is in shape similar to its final saturation behaviour). Not only does this assumption realize the simplest model, it is essential for Verhulst's equation as the algebraic expression of the failed change model. It is, of course, an empirical issue whether this presupposed symmetry of the peak is legitimate or not. The data suggest it is tenable, though a 4-parameter model, obviously, gives better fits than a 2-parameter model. The symmetry assumption is natural in a theory that models failed changes in a principled way.

Tony Kroch (*pers. comm.*) and an anonymous reviewer suggest giving an independent interpretation to the two S-curves, S_1 and S_2 , that constitute the failed change according to the formula $F = S_1 \cdot (1-S_2)$. The factor S_1 , describing the rising flank of the failed peak, might then be connected to some (successful) change in one direction, and S_2 , describing the declining flank of the failed peak, might be connected to another (successful) change in the other direction. This is in fact the accidental failure model. Under such a model, there is no principled reason to assume both slope parameters that jointly constitute the peak to be equal. Moreover, there would be no reason that the two actuation parameters should be equal either, the only restriction being that the actuation of S_2 is after S_1 . Let us apply this suggestion to *do*-support, where the failed change (positive declarative *do*, the lowest solid curve in Figure 3) is tied to successful changes in negative and interrogative contexts (especially the initial dashed curve in Figure 3). In terms of the features [s] and [f] of Section 4.2, this proposal can be formulated as follows. The rise and decline of *do*-support can be analysed as [-s] = V-to-I and [+f] = positive declarative *do*-support (PAD). The PAD-peak can then be analysed as in (10), where PAD is split into two periods, one in which [+s, -f] competes with verb raising [-s] (this is period 2) and one in which [+s, +f] competes with affix hopping [+s, -f] (this is period 3), both following an S-curve. The assumption is now that these two S-curves, multiplied with each other, form a peak.

$$(10) \quad \begin{array}{cccc} \text{period 1} & & \text{period 2} & & \text{period 3} & & \text{period 4} \\ [-s] & \rightarrow & [-s] & \rightarrow & [+s, -f] & \rightarrow & [+s, -f] \\ & & [+s, +f] & & [+s, +f] & & \end{array}$$

Now, as one reviewer argues, since PAD (i.e. [+s, +f]) competes with different alternatives in period 2 and 3, (with V-to-I and affix hopping, respectively), there is no reason why the CRH would predict an equal slope parameter to the rising logist S_1 and the declining logist S_2 . This approach, therefore, allows for an *asymmetrical* peak for the failed change, its increase tied to the slope parameters of the other *do*-support processes (the S-curves of negative *do*-support), its decline to some other process. Now, the only slope parameter that governs the change is the slope parameter of the S-curves, which are, according to the CRH, all connected to each other. So there is no possible slope parameter within the overall process. So we either state that the slope

14. An alternative mathematical representation is: $\frac{1}{2 + 2\cosh((t-t_0)/a)}$

parameter of the peak's right flank is equal to the slope parameter of its left flank, or we find a second process to which it is tied. The onus of finding this process is on this asymmetric model approach, which is probably at odds with the CRH itself. For, if such process could be found at all, it must be triggered by a parameter switch that is different from the one we are investigating, and *do*-support as a unitary process would be lost. In this study we simply adopt the null hypothesis, that the slope parameter of the decline is equal to all the other S-curves, in full agreement within the CRH. This is, of course, an empirical question. So, if these slopes turn out to be systematically and significantly unequal (with opposite sign), our model would be incorrect. Unfortunately, the test must be left for further research as none of the data are sufficiently detailed to perform this discriminating test.¹⁵

4.4 Explaining the model

In this section we discuss the two ingredients of the mathematical model sketched in the previous section. The first ingredient is the assumption that failed changes exist in natural language. This raises the question: how can a change be an inherent failure? The second ingredient is the hypothesis of a flanking successful change that is identical to the constituting logistic function within the failed change. This raises the question: how can a failed change be dynamically connected to a successful change? In order to legitimate these assumptions, we must return to the nature of linguistic innovations. The older assumption in the principles and parameters model of generative grammar is that innovations occur as an imperfect transmission in L1 learning. Language change is a change of the grammar in L1 learners (Lightfoot 1999, Andersen & Lightfoot 2002). The reanalysis model is the standard implementation of this idea. This L1-scenario has faced various objections: why would language change? Are children such bad learners? It conflicts with the idea that the language acquisition device is extremely robust (Pinker 1984). And there is, of course, the conflicting data problem: if children have a grammar distinct from their parents, how do they deal with the conflicting input data that will certainly be heard?

To avoid some of these problems, it has been proposed that L2 acquisition be the locus of language change (Bybee & Slobin 1982; Aitchison 1991). These authors locate the change in peer group innovations, which subsequently spread through the language through social networks. Now, the status of L2 learning has been an issue of much research and debate (cf. Epstein, Flynn & Martohardjono 1996). Despite the differences in approach, it is commonly felt that the essential difference between L1

and L2 learning is that L2 learning involves persons after the maturation period, that is, persons who do not have full access to Universal Grammar (UG) or the Language Acquisition Device (LAD) and cannot acquire new parameter settings after the maturation. It may be clear that the idea of adult innovations as the trigger of language change poses a problem for the parameter theory, as it must be assumed that adults are able to set (new) parameters. Consequently, the L2-hypothesis cannot be applied to the parameter model of grammar. Some enriched model is necessary.

An interesting model of language change where both L1 and L2 factors are active is found in Weerman (1993). Basing himself on Clahsen & Muysken (1986), Weerman formulates two principles of acquisition:

- (11) a. Children cannot deviate from positive evidence
- b. Parameter change can only take place under L1 acquisition

The first principle is taken from a deterministic hypothesis of acquisition. The second is the standard maturation assumption in acquisition studies. The two principles exclude change through parameter resetting in the Principles & Parameters model, because, firstly, children cannot deviate from what is offered, and secondly, only children can set parameters. To escape the dilemma of a deterministic connection between E-language exposure and acquired settings in the I-language, Weerman suggests a 'peripheral rules solution'. This model adds a secondary ingredient to the model that determines the interaction between E and I-language. Apart from standard L1 acquisition by setting parameters, Weerman assumes *adult innovations* which are stored in the I-language in the form of *peripheral rules*. Peripheral rules are rules that do not belong to the core grammar but are applied to the output strings. Application of peripheral rules is not produced by parameterized UG, but by later sociological accommodation. These peripheral rules are active in later stages of first language acquisition ('late L1') as well as in the process of second language acquisition. Reformulated in modern terms, peripheral rules are not produced by the language faculty proper but by the human problem-solving faculty and, hence, subject to variation. In a system enriched with peripheral rules we can model language change under retention of the principles in (11). Weerman's peripheral rules solution to language change boils down to the following sequence of events: parameter settings are transmitted flawlessly to the next generation, but adolescents and adults change their output by adding peripheral rules (they cannot change parameter settings). The next generation, now, taking all data as primary linguistic data, construct a new parameter setting based on the complete input. Speaking on the change from OV to VO in English, Weerman writes about the increased number of VO-patterns:

"My suggestion is that these (VO) leakages were so to speak exaggerated via L2-M acquisition [adult innovations; GJP], both quantitatively and qualitatively. The

15. The rather technical reason is that the fitting procedure becomes unstable in this asymmetric model. Being de facto a 5-parameter model, it is extremely flexible and allows for too many, often bizarre, best fits.

relevant speakers do not change their internalised setting of the head parameter. From the perspective of their L1-grammar these over-generalisations are *ungrammatical* [my italics; GJP]. What they do is add a peripheral rule. A next generation, however, could set the head parameter". (Weerman 1993).

We will speak of 'off-grammatical' constructions whenever innovations are not produced by core grammar but by peripheral rules.

A similar, though less explicit, proposal is found in Van der Wurff (1990: 30). Van der Wurff considers a language produced by grammar G2 to which isolated new constructions (which he calls C_c , be they borrowings or innovations from G1) are attached. If these innovations are compatible with G2, basically nothing happens. Suppose, however, the more interesting case that these additional constructions are incompatible with the grammar G2. This creates an unstable complex (G2, C_c) which is then transformed to a stable (G3, C_c). In general, language change proceeds through such unstable intermediate complexes.

A more interface-type approach of the adaptive nature of L2 acquirers can be found in Vainikka & Young-Scholten (1994). They show that L2 learners fall short on morphology or, rather, they fall short on the fine tuning between syntax and morphology, that is, the alignment of morphological forms and the accompanying syntax. An example of such a mismatch between morphology and syntax is an interlanguage that has lost rich inflection but retains V-to-I or vice versa. The idea of failed changes makes it possible to operationalize these ideas.

Combining the unstable intermediate complexes with the morphology-syntax misalignment proposed by Vainikka and Young-Scholten with Weerman's off-grammatical adult innovation, we may think of a failed change as an off-grammatical adult innovation that serves some linguistic or sociological goal but does not have the perfect alignment of morphology and the parameter settings. A successful change, on the other hand, is a change that is compatible with all principles of UG and has a perfect alignment. Successful changes are introduced by L1 learners, for instance as accommodations of the failed change initiated by adults. Let us, therefore, represent language as a pair consisting of a parameterized Grammar G1 and a lexicon D1. Suppose adults start using a new lexeme (be it a deflected verb or a pronoun with a different specification, etc.), that is, their lexicon changes into D2 that is not fully compatible with the old grammar. They fail to align the lexeme and the grammar properly: *[G1, D2]. Alternatively, they can change the grammar but fail to align it properly with the morphology: *[G2, D1]. Both situations are unstable and resolve to a fully aligned [G2, D2]. Only non adult speakers are able to make this alignment before the maturation time. We call it an L1-accommodation of the instability represented by [G1, D2] or [G2, D1].

(12)	Old situation	→ Intermediate situation	→ New situation
	grammatical	off-grammatical mismatch adult innovations (L2)	grammatical L1-accommodations
	[G1, D1]	*[G1, D2] or *[G2, D1] failure	[G2, D2] success

This scenario, essentially taken from Weerman (1993) and Van der Wurff (1990), resolves various problems stated above, most notably the relation between the failed and the successful change. The problem of the conflicting data in the reanalysis scenario disappears as well: these conflicting data are the very *trigger* of the change. The essential ingredient of the model is that it allows for off-grammatical data, that is, data without full alignment between syntax and morphology for a relatively short time: the innovation inherently fails, but may have an impact on the system.

4.5 Applying off-grammaticality to the case study

In order to illustrate the proposed L2 mechanism of language change and the role of failed changes in it, let us apply it to our case study of the rise of the reflexive in Dutch. First we must adopt the assumption that *sick* was first introduced into Dutch to solve the growing unsuitability of *hem* as a reflexive. The unsuitability of *hem* was felt with the L1 speakers of Drentish. (For the reason why *hem* became unsuitable, cf. Section 3.2 and Postma 2004). So, the adult speaker borrowed *sick* or *sich* into their language. Both *sick* and *sich* were suitable as reflexives since they display the desirable number neutralization. Obviously, this suitability is not accessible to the innovative adults, they simply borrowed these upon the unsuitability of *hem*, but the suitability of *sick/sich* is accessible to the next generation. Though feasible as reflexives, both *sick* and *sich* were as such incompatible with the structure of the Dutch morphological system (to be elaborated below), and should fail unless something happened. The idea is that this something was possible with *sich* but not with *sick*. Let us start with the incompatibility. The reason why *sick/sich* are incompatible can be extracted from the scheme in (13).

(13) Accusative reflexive pronouns

	Drenthe	new	Low German	Central German
1p	mi	mi	mik	mich
2p	di	di	dik	dich
3p	hem	*syk/sych	sik	sich

In (13) we give the (reflexive) pronominal system ('me/myself, you/yourself, him(self)') in Drenthe of 1400 and of two types of German. The German pronouns have a transparent accusative marker *-k/ch*. After specialisation for number, *hem* was replaced by

inclusion of failed changes confirms the superiority of the logistic model on the basis of the additional qualities listed in (16).

- (16) – The logistic is a solution of a simple differential equation
 – The model is applicable to failed changes
 – The model relates failed and successful changes algebraically
 – The failed change is the first derivative of the successful change
 – The successful change is the cumulative function of the failed change
 – The model allows ephemeral virtual changes (in the E-language but outside the I-language)

Furthermore, the failed change model has the theoretical interpretation given in (17).

- (17) – Failed changes are the initiators of successful changes
 – Failed changes are off-grammatical adult/L2 innovations
 – Successful changes are L1-fixings of off-grammatical adult/L2 innovations

Despite the advantages, there are some problems with this model. These show up when we try to apply the model to the rise of *do*-support in the history of English.

6. *Do*-support in positive declaratives in English revisited

In the previous sections we designed an algebraic model that included failed changes along with successful changes. The failed change is proportional to the first derivative of the successful change. The logistic model constitutes a two-parameter model with the actuation time and the slope as its parameters. These two parameters define when and how quickly a change proceeds from a penetration of 0 (or 0%) to a penetration of 1 (or 100%). Notice that in this model the occurrence rate of the innovation is normalized with respect to the absolute occurrence of the construction. It is therefore strictly speaking a 3-parameter model. If the failed and the successful innovations occur in exactly the same pre-defined set of constructions, that is, when they target the same 'market', the model predicts that the failed change will reach a maximal penetration of 1/4 (25%) before it declines. This is a reasonably correct prediction for the rise of *sick* in Dutch. In most of the cases, however, there is no predefined target. In brief, there is a normalization problem in the modelling of successful changes, which becomes even more urgent in the modelling of failed changes (we cannot simply choose some relatively late *t*). Normalization is especially difficult in its relation to the successful change if the failed change targets another set of constructions. The normalization problem is one of the reasons that we find deviations from the predicted 25%: for example, it is much too low in the failed *do*-support (8%), displayed in Figure 4, repeated here as Figure 11.

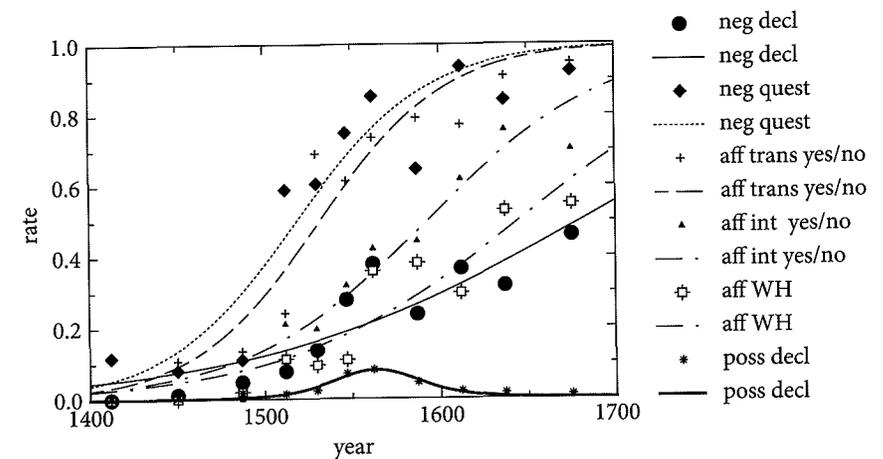


Figure 11. The rise of *do*-support in English on a linear scale (data from Ellegård 1953)

Why is the failed change rate so low in some cases? In this section we address the normalization problem for the case of the failed *do*-support, as it has been discussed from the very beginning and was one of the reasons for Ellegård to start a quantitative analysis. Therefore, *do*-support is illustrative for what problems the analysis may encounter with respect to normalization.

6.2 The normalization problem

The earlier scholars who studied *do*-support noticed that the occurrence of periphrastic 'do' was initially (i.e. during the 15th century) predominantly observed in positive declarative clauses (Dietze 1895; Trnka 1930: 53; Engblom 1938), that is, precisely where it no longer occurs in modern English. These scholars came to this impression because they considered the set of *do*-support clauses. In this set, the majority of clauses are positive declarative. However, as Ellegård noticed, to draw a conclusion of dominance in general from this fact is a methodological flaw as positive assertions exceed by far the number of interrogative and negative sentences (by roughly a factor 17). Ellegård's quantitative work made it possible to shift from considering absolute frequencies to relative frequencies in the respective constructions. Ellegård argues that there is "absolutely no justification for supposing that the frequency was at any time higher in affirmative sentences than in others". Ever since, the rate numbers have been calculated proportional to the occurrence rate of the construction in which they show up. All data in Kroch (1989) are normalized in this way. From the perspective of I-language this seems justifiable. Things change, however, if we calculate the impact of a change in a construction (for instance a failure) on other constructions (where it

might become successful). This might be an E-language phenomenon and hence the absolute rates might be important.

The nature of the too small peak of the failed change in the case of *do*-support (8%) can be understood from the perspective of the overall increase of *do*-support. The relevant data are not in Kroch (1989) but can be found in Ellegård's Table 7. We have plotted these in Figure 12 (the dots).

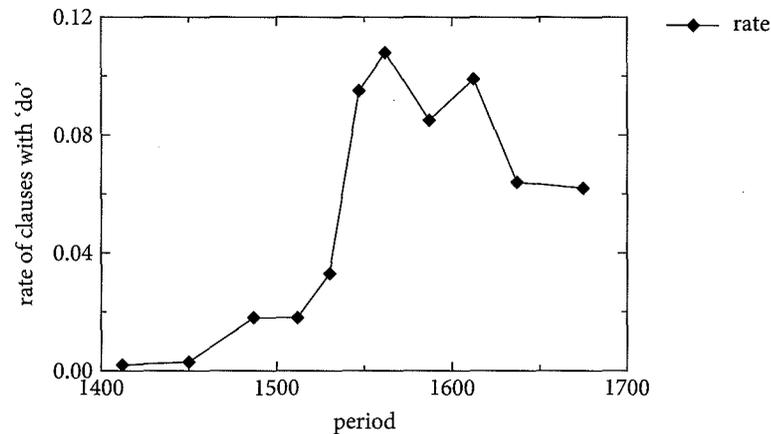


Figure 12. Overall rate of the *do*-construction (data from Ellegård 1953: table 7)

We see that *do*-support starts out in the early fifteenth century, increases rapidly in the sixteenth century, declines slightly and goes to a saturation of 7% during the seventeenth century. Interestingly, the increase displays an overshoot up to 11% around 1550, the nature of which is not discussed in Ellegård, nor in the literature afterwards. From the 'failed change' perspective, it is natural to identify the peak with the failed change, whose peak coincides with the inflection point of the successful change. The question is of course whether the failed change perspective can appropriately model these data.

The modelling of the empirical data given by Ellegård needs requires careful scrutiny. As said, the increase in *do*-support has a peak around 1550 and then goes asymptotically down to the level of 7%. This value is the relative number of negative and interrogative contexts in corpora in general. This reflects the fact that the rate of *do*-support in Neg and Q contexts goes to 100%. In contrast to the case of *sick/sich*, the failed change in the case of *do*-support is not part of the successful change but realizes itself *outside* the domain of success. The failures are not *included* in the success, but cause a *peak that is superposed to it*. This peak is rather strong because of the fact that the contexts of affirmative declaratives are much more numerous than the contexts of successful changes (in negative and interrogative clauses). In this case we need two

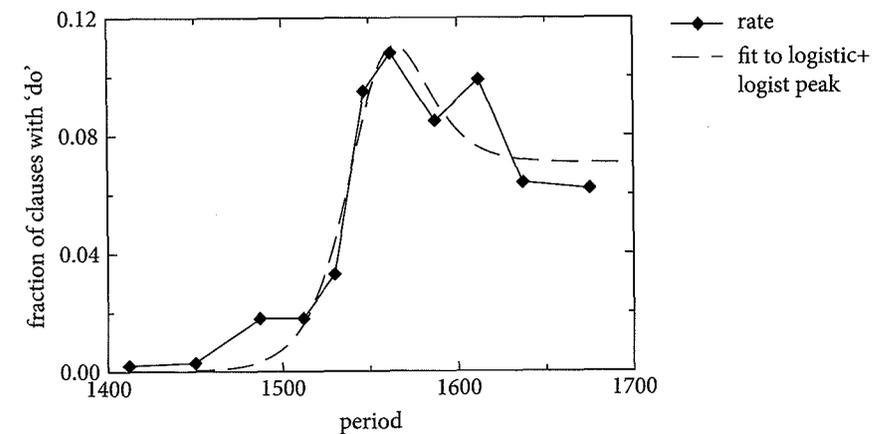
normalization factors, a normalization factor of the overall process (a_3) and a relative weight factor of the failed and the successful change (a_4). So we arrive at the formula in (18) (For a derivation, see the appendix).

$$(18) \quad \text{Rate} = a_3 \cdot [S + a_4 \cdot F], \text{ with } S \text{ a logistic function and } F \text{ a logistic peak}$$

As to the internal parameters (a_1, a_2) of S and F , these are as follows. According to our hypothesis, the activation and the slope of the successful and failed changes are equal. So we predict that the measured data should fit to a function of the form in (19).

$$(19) \quad \text{Rate} = a_3 \cdot [S(a_1, a_2) + a_4 F(a_1, a_2)]$$

where a_1, a_2 are the activation time and the slope parameter. In this fit, the slope and the activation parameter of the failed change (the peak) are taken to be identical. The third parameter, a_3 , is the saturation level, that is, the proportion of negative and interrogative constructions that are feasible for *do*-support ($7\% \pm 2$). This parameter is needed in any growth process (already for dimensional purposes). The only additional parameter in this model is a_4 , the relative strength of the peak and the logistic function. In Figure 13 we show a best fit of this function to the data given by Ellegård. It shows that this model is quite faithful to the data.



$$\text{Chi squared} = 1.1473e-3$$

Parameters:	Standard deviations:
a[1] = 1557.2296	$\Delta a[1] = 5.9884$
a[2] = 15.0347	$\Delta a[2] = 3.2376$
a[3] = 7.0679e-2	$\Delta a[3] = 8.5145e-3$
a[4] = 4.0024	$\Delta a[4] = 1.1146$

Figure 13. Overall rate of *do*-support with algebraic fit in the failed change model

We conclude that the failed change model can be adequately applied not only to cases where the failed change is part of the successful change but also to cases where the failed and the successful change apply in separate syntactic contexts.

6.3 Calculation of the 8% level of the failed change

Now that we have disentangled the failed and the successful change algebraically, we are in a position to normalize the failed change peak. Consider the decomposition of the curve in Figure 14 into a successful part and a failed part.

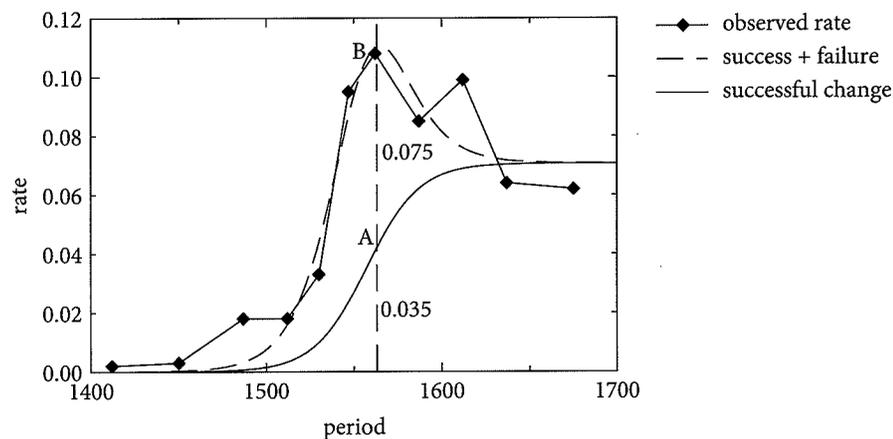


Figure 14. Data analysis of *do*-support in the failed change model

The peak of the failed changes corresponds with the inflection point of the successful change. Now, the failed contribution in the overall rate of *do*-support is the segment A–B, which is 0.075, the level at the peak value A (0.11) minus the level at inflection point B of the successful change, 0.035 (= 0.07/2). In order to get the rate of *do*-support in positive declarative clauses, we must not normalize it to the total number of clauses but to the positive declarative clauses only, that is, 100%–7%, the ratio of WH and Neg clauses. Using the subtraction $1 - 0.07 = 0.93$, we arrive at the value of 8% for the peak of the failed change, according to the calculation in (20).

$$(20) \quad 0.11 - 0.035 / 0.93 = 0.075 / 0.93 = 0.08$$

This corresponds to the measured value of 8% in Figure 11. We conclude that the low peak of the failed change in the relative frequency diagram in Figure 11 is a mere consequence of the numerical dominance of positive declarative clauses. If we use absolute levels of occurrence, both the successful and failed changes have the same normalization factor. There is an important theoretical implication of this result. It means that the failed and successful change interact at absolute occurrence levels, not

at relative occurrence levels. Put differently, the interaction of failed and successful change occurs in the domain of E-language, not at the level of I-language. The relative occurrence level, on the other hand, is a consequence of the high occurrence of positive declarative clauses (where it fails) and the low occurrence of negative and interrogative (where it succeeds).

7. Should all linguistic data be considered grammatical?

While all linguistic data are to be taken with the same scrutiny, it is far from clear whether all data are to be considered grammatical. Especially in periods of linguistic change, the general assumption that all data are grammatical cannot be maintained, for, if this were so, no change would be possible. One way to go is to assume a two-grammar system where the data that are ungrammatical in one system are taken to be grammatical in the other and vice versa. This approach is generally applied in generative studies and captures the change as a gradual diffusing replacement of one grammar by the other in a community by direct grammar competition (Kroch 1989; Santorini 1992; Henry 1995; Pintzuk 2002; Postma 2004). This grammatical approach considers the extent to which the two grammars overlap with respect to their respective outputs to be the possible locus of (the) change and has functioned as a successful heuristic tool in diachronic research. Within this overlap, reanalysis is possible: a string produced by one grammar can be parsed within the other. However, the reanalysis hypothesis, with Lightfoot as the major advocate, has struggled with the conflicting data problem: on what basis does an L1 learner ignore data that conflict with his new parameter setting? As yet, this problem has not been overcome. Furthermore, the L1 approach does not give consideration to the influence of small sociolinguistic changes under influence of peer groups who deliberately change their E-language (Labov 1972; Bybee & Slobin 1981; Sankoff & Blondeau 2007).

Another way to go is to assume that linguistic change proceeds through a linguistic stage that includes fundamentally off-grammatical data. Under the assumption that L1 acquisition fully complies with the rules of UG and produces configurations that are generated by an instance of UG, the off-grammatical approach utilizes L2 changes imposed by adult speakers or adult L2 learners. Such changes imposed by L2 learners, then, (may) cause a reshuffle in the core grammatical system. By its very assumptions, the second approach deals with the relation between E-language and I-language, and is fundamentally sociolinguistic. The obvious problem here is how to distinguish the grammatical and the off-grammatical data.

In this study, we have developed a simple model that is a variant of Kroch's model and which includes the sociolinguistic factors that feed the transition as claimed by Labov. We have argued that to any change that follows an S-curve (the so-called 'successful change'), there is a 'failed change' of sociolinguistic nature that

fuels the successful change. The unsuccessful change consists of off-grammatical variants that die out after a while. In those cases where we can empirically distinguish the off-grammatical data from the grammatical data, the peak-type curve of the unsuccessful change maintains an algebraic relation with the successful change. In this way, we take the transitional data, despite their off-grammaticality, seriously as produced by peripheral rules. More precisely, we have shown that these transitional data are off-grammatical in the sense that they are ungrammatical in the system both before and after the linguistic change. Nevertheless, these off-grammatical changes do not escape from Kroch's (1989) Constant Rate Hypothesis.

Appendix

If we ignore the number of focussed positive declarative *do* sentences (just as Ellegård does), we have

$$N(do) = Suc(do) + Fail(do)$$

Hence, dividing both members by TOT, the total number of clauses considered, we get:

$$\frac{N(do)}{TOT} = \frac{Suc(do)}{TOT} + \frac{Fail(do)}{TOT}$$

Furthermore, if we divide and multiply by NPI (the total number of WH and NEG clauses) and divide and multiply by POS (the total number of positive declarative clauses), we obtain:

$$\begin{aligned} \frac{N(do)}{TOT} &= \frac{Suc(do)}{TOT} + \frac{Fail(do)}{TOT} = \frac{Suc(do)}{NPI} \cdot \frac{NPI}{TOT} + \frac{Fail(do)}{POS} \cdot \frac{POS}{NPI} \cdot \frac{NPI}{TOT} \\ &= \frac{NPI}{TOT} \left[\frac{Suc(do)}{NPI} + \frac{Fail(do)}{POS} \cdot \frac{POS}{NPI} \right] \\ &= a_3[S + F \cdot a_4] = \end{aligned}$$

In sum,

$$rate(do) = a_3[S + F \cdot a_4]$$

where, a_3 is the overall fraction of interrogative and negative clauses (NPI/TOT) and a_4 is the overall fraction of positive declarative and interrogative + negative clauses (POS/NPI). These ratios are not known, but are supposed to be constant in time and can be treated as fitting parameters to the graph. Results are $a_3 = 0.07$ (7%) and $a_4 = 4.0$. Note that we have taken the rate of *do*-sentences with respect to all clauses, modal clauses included.

References

- Aitchison, Jean. 1991. *Language Change: Progress or Decay?* Cambridge: CUP.
- Aldrich, John Herbert & Nelson, Forrest D. 1984. *Linear Probability, Logit and Profit Models*. Beverly Hills CA: Sage.
- Altmann, Gabriel, Buttlar, Haro von, Rott, Walter & Strauß, Udo. 1983. A law of change in language. In *Historical linguistics* [Quantitative Linguistics 18], Barron Brainerd (ed.), 104–115. Bochum: Brockmeyer.
- Andersen, Stephen & Lightfoot, David. 2002. *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge: CUP.
- Bailey, Charles-James. 1973. *Variation and Linguistic Theory*. Washington DC: Center for Applied Linguistics.
- Barbiers, Sjeff & Bennis, Hans. 2003. Reflexives in dialects of Dutch. In *Germania et Alia: A Linguistic Webschrift for Hans den Besten*, Jan Koster & Henk van Riemsdijk (eds). (<http://www.let.rug.nl/~koster/DenBesten/contents.htm>).
- Bybee, Joan & Slobin, Dan. 1982. Why small children cannot change language on their own – evidence from the past tense. In *Papers from the Fifth International Conference on Historical Linguistics, Galway, April 6–10 1981*, [Current Issues in Linguistic Theory 21], Anders Ahlqvist (ed.), 29–37. Amsterdam: John Benjamins.
- Clahsen, Harald & Muysken, Pieter. 1989. The availability of universal grammar in adult and child learners: A study of the acquisition of German word order. *Second Language Research* 2: 93–119.
- Dietze, H. 1895. *Das umschreibende Do in der neuenglischen Prosa*. Ph.D. dissertation, Universität Jena.
- Ellegård, Alvar. 1953. *The Auxiliary do: The Establishment and Regulation of its Use in English*. [Gothenburg Studies in English]. Stockholm: Almqvist and Wiksell.
- Engblom, V. 1938. *On the Origin and Early Development of the Auxiliary do* [Lund studies in English 6]. Lund: C.W.K. Gleerup.
- Epstein, Samuel, Flynn, Suzanne & Martohardjono, Gita. 1996. Second language acquisition: Theoretical and experimental issues in contemporary research. *Brain and Behavioural Sciences* 19: 677–714.
- Fontaine, Carmen. 1985. Application de methodes quantitatives en diachronie: L'inversion du sujet en français. MA thesis, Université du Québec à Montréal.
- Henry, Allison. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: OUP.
- Hinterhölzl, R. 2004. Language change versus grammar change: What diachronic data reveal about the distinction between core grammar and periphery. In *Diachronic Clues to Synchronic Grammar*, Eric Fuss & Carola Trips (eds), 131–160. Amsterdam: John Benjamins.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199–244.
- Labov, William. 1972. *Language in the Inner City: Studies in Black English Vernacular*. Philadelphia PA: University of Pennsylvania Press.
- Lightfoot, David W. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.
- Naro, Anthony J. 1984. Review of 'Historical Linguistics' by Barron Brainerd. *Language* 60: 929–932.

- Noble, Shawn. 1985. To have and have got. Paper presented at N.W.A.V.E. 14, Georgetown University.
- Oliveira e Silva, Giselle. 1982. Estudo da regularidade na variação dos possessivos no Português do Rio de Janeiro. Ph.D. dissertation, Universidade Federal do Rio de Janeiro.
- Pinker, Steven. 1984. *Language Learnability and Language Development*. Cambridge MA: Harvard University Press.
- Pintzuk, Susan. 2002. Verb-object order in Old English: Variation as grammatical competition. In *Syntactic Effects of Morphological Change*, David Lightfoot (ed.), 276–299. Oxford: OUP.
- Piotrowskaja, A.A., & Piotrowskij, R.G. 1974. Matematičeskie modeli diachronii i tekstoobrazovanija. In *Statistika reci i avtomatičeskij analiz teksta*, 361–400. Leningrad: Nauka.
- Postma, Gertjan. 2004. Structurele tendensen in de opkomst van het reflexief pronomen 'zich' in het 15de-eeuwse Drenthe en de Theorie van Reflexiviteit. *Nederlandse Taalkunde* 9: 144–168.
- Priestley, L. 1955. Reprise constructions in French. *Archivum Linguisticum* 7: 1–28.
- Reinhart, Tanya & Reuland, Eric J. 1993. Reflexivity. *Linguistic Inquiry* 24: 657–720.
- Sankoff, Gillian & Blondeau, Helene. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83(3): 560–88.
- Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory* 10: 595–640.
- Trnka, B. 1930. *On the Syntax of the English Verb from Caxton to Dryden* [Travaux du Cercle Linguistique de Prague 3]. Prague.
- Vainikka, Anna & Young-Scholten, Martha. 1994. Direct access to X' structure: Evidence from Korean and Turkish adults learning German. In *Language Acquisition Studies in Generative Grammar* [Language Acquisition & Language Disorders 8], Teun Hoekstra & Bonnie Schwartz (eds), 265–316. Amsterdam: John Benjamins.
- Verhulst, Pierre-François. 1838. Notice sur la loi que la population poursuit dans son accroissement. *Correspondance mathématique et physique* 10: 113–121.
- Weerman, Fred. 1993. The diachronic consequences of first and second language acquisition: The change from OV to VO. *Linguistics* 31: 903–931.
- Weinreich, Uriel, Labov, William & Herzog, Marvin I. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics: A Symposium*, Winifred P. Lehmann (ed.), 95–195. Austin TX: University of Texas Press.
- van der Wurff, Wim. 1990. Diffusion and Reanalysis in Syntax. Ph.D. dissertation, Leiden University.