

Literary debutants and their career:
Event history analysis of author's reputation

NEL VAN DIJK

Department of Language and Literature

Tilburg University

JEROEN K. VERMUNT

Department of Methodology

Tilburg University

Corresponding author:
Nel van Dijk
Department of Language and Literature
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
E-mail: n.vandijk@kub.nl

Literary debutants and their career: Event history analysis of author's reputation

NEL VAN DIJK and JEROEN VERMUNT

Abstract

The subject of this article is literary authors' reputation. We are interested in the factors underlying stability and instability in literary careers. This topic is only one aspect of a larger research project that is devoted to institutional resources and constraints affecting a writer's career.¹ This article pursues to improve our current knowledge of what creates an authors' reputation by empirical research. Besides, our aim is to illustrate the applicability of event history analysis, a methodology for the analysis of longitudinal data. Different kinds of cultural products have a life cycle of their own; for instance, over time authors may be subject to an increase or a decrease in reputation. Event history analysis is a technique that makes it possible to model such dynamic processes.

1. Introduction

Literary careers develop very differently. Besides the (relatively few) authors whose books are always serious candidates for the top-10 of most read books, there are many writers whose new works are hardly noticed by the literary critics and the reading public.² These differences can not exclusively be accounted for by looking at intrinsic qualities of literary works. The valuation of literary works is limited in time and place. In different periods, opinions on the literary value of cultural objects vary. Sometimes highly esteemed works fall into oblivion, while others are reassessed. Even at one point in time different judgments on a literary work can co-exist.

The major dissimilarities between literary careers are partially the result of characteristics of the profession. First, authorship is a free profession; one does not need a diplom to become a writer. There are no strict conditions or rules a debutant author has to comply with. Second, only a very small percentage of all writers is capable of making a living out of writing fiction. In the Netherlands, most authors have to make use of the possibility to sign in for a so-called working-grant with the 'Foundation for Literature' a national organization that supports writers. Furthermore, many writers perform other activities in order to get a source of income. Those activities can vary from holding lectures on literary subjects till working within a totally different sector such as education.

Authors may also be more or less successful because of factors that are related to the reception of their work. Each new title can be seen as an event in time, that receives more, less or an equal amount of attention in the literary field. In our opinion, the dissimilarities between literary careers can also be explained by factors that are non-textual, and that are related to this process of assigning quality to literary works. The questions posed in this article deal with the factors that influence the careers of writers. We want to determine which non-textual aspects are important for a writer in order to operate in a successful way.

For non-literary careers income usually is the most important indicator of success. That is, however, not true for literary careers. As already mentioned, only very few writers are able to make a living out of writing fiction. In addition, the relation between art and commerce has always been rather difficult. Financial profit and a serious performance of a creative job do not go together (Bourdieu 1983). When we call an author successful, we more often refer to the reputation of this author within the literary field than to his or her income. This reputation can be made clear by looking at the critical attention a literary work receives. The descriptions and judgments in reviews influence both the reading public as well as other critics in the literary field. The continuity and broadness of the attention paid to a body of work by dailies and weeklies, together with the way that attention is being articulated, can be seen as an important indicator of the value that is attached to an author and his or her body of work. In investigating the role of criticism in gaining a reputation as a writer, both qualitative and quantitative modes of research are needed to specify the factors involved.³ Here we would like to report on the results of a quantitative research.

2. Debutant's data

On an earlier occasion Van Rees and Vermunt (1996)⁴ reported on their research after literary authors' reputation. They were the first to investigate the use of event history analysis for this purpose. The present research is, however, based on a larger data-set. That makes it possible to measure the dependent variable in a more refined way and to use a larger set of independent variables. We aimed at accumulating our present knowledge by formulating hypotheses with the results from the earlier research in mind.

We performed a quantitative analysis on Dutch authors' data derived from a survey of the critical reception of 53 writers. Selection criteria were: authors made their debut in prose or poetry in the period from 1975 till 1979. So, we could follow them over a long period of time, about 20 years. We collected data on all titles (468 in total) these 53 authors produced between 1975 and 1994. Some 1750 reviews were examined. Our data-set differs from that of Van Rees and Vermunt (1996) on some points. They used the number of reviews in top-periodicals as the measure for the amount of attention. We measured attention as the number of reviews in national dailies and weeklies. In the Dutch situation, with a population of 15,5 million, most national newspapers are made in 'the Randstad', a conglomerate of several big cities. Those newspapers are also read in other parts of the country. Furthermore, different areas in the country have their own regional newspapers. Van Rees and Vermunt selected 18 authors who made their debut with a prose title. Furthermore, with regard to the titles following the debut-title, they collected reviews on prose-titles only, whereas we collected data on all titles, irrespective of the genre. The size of their data-set is not much smaller than ours. This is caused by the fact that prose titles receive more reviews than the other genres. So we have more authors and titles, but a smaller average number of reviews per title.

A strong point of the data set on the careers of 53 Dutch debutant authors is that it is of a longitudinal

nature: for each author there is information on the reviews of all their titles which appeared during a long period of time. This means that information is available on the change in attention between subsequent titles. That is the reason why the analysis is focused on explaining differences in changes in attention instead of differences in attention. The central question to be answered can be formulated as follows: *Which factors determine whether an author experiences an increase, a decrease, or no change in attention between subsequent titles?* To answer this question it is necessary to define the basic time dimension which is used in the analysis, to select the most appropriate measure for the dependent variable 'amount of critical attention', and to specify the explanatory variables to be used in the model.

3. Time dimensions

The time dimensions that are most frequently used in event history analysis are calendar time, process time, and waiting time. In the case of the analysis of changes in attention, these time dimensions can be equated to the month or year in which a title is published, the time since the start of an authors' career, and the time since the previous change in attention, respectively. In the current analysis, however, another time dimension seems to be more appropriate, namely, the rank order of the title. The main reason for using this discrete variable as time dimension, is that changes in critical attention can only take place when a new title comes out.

When the rank order of a title is treated as the time dimension, the authors' data can also be seen as discrete-time event history data. This means that for each point in time, that is, for each title, there is information on the dependent and independent variables of interest. Therefore, a discrete-time event history model was used to specify a regression model for changes in attention. To be able to use such a model, the data have to be organized into a so-called person-period file. That is, a file with titles as records, rather than with reviews or authors as records (Allison 1982; Yamaguchi 1991). It should be noted that since we defined the time variable as the rank order of a title, the length of the observation period in terms of number of titles is not equal for each author. Some authors publish a vast body of work over a couple of decades, while others publish only a few books in the same period. This, however, does not complicate the research. Event history analysis makes it possible to deal with observations of different length.

The 53 authors in the data set produced 468 titles. But since the first title cannot be used to model transitions between titles (the first title cannot be compared to a previous title, so we cannot speak about an increase or a decrease of attention), a file was constructed with 415 records. The number of titles per author after the debut ranged between 1 and 24.

4. Defining changes in reviewers' attention

The first step in performing event history analysis is defining the events of interest. In our case, an event is a change in 'amount of attention', the variable we want to explain. As mentioned above, in the present analysis, the number of reviews in national dailies and weeklies was used as the measure for the amount

of attention. To make an event history analysis feasible, the dependent variable 'amount of attention' was collapsed into five categories. These categories which are shown in the top row of table I are: 0, 1-2, 3-4, 5-6, and >6 reviews. They represent the states that an author can occupy at each 'point in time'.

[INSERT TABLE I ABOUT HERE]

An event is defined as a *transition from one state to another state between subsequent titles*. For instance, the transition from one or two reviews of the second title to five or six reviews of the third title. An impression of the mean transition probabilities among the five different states can be gained from table I. This table reports the transition probabilities which are obtained if one assumes them to be independent of the rank order of the title and the explanatory variables we used in the analysis. For instance, it can be seen that an author with 0 reviews of a title has a probability of 0.25 of getting 1-2 reviews of the next title ($P[1-2|0]$) and a probability of 0.20 of seeing the number of reviews increased to 3-4 ($P[3-4|0]$). Note that for an author who was in the lowest or in the top-category with the previous title, there is a high probability that he or she will remain in the same category with the next title (0.45 and 0.44, respectively).

Note, furthermore, that it is not very likely that an author improves or deteriorates more than two levels of attention between two titles: the values of $P[0|5-6]$ and $P[0|>6]$ are very low (0.04 and 0.06), so is the value of $P[>6|0]$ (0.07). In general, we can say that the probability of having an event is high, which shows that this part of the literary field is quite dynamic. Nevertheless, most events represent little movements rather than large ones. Reviewing is a conservative process, large fluctuations in the attention paid to literary works are not very likely.

A problem with the transition probabilities appearing in table I is that there are many of them; there are 20 independent probabilities. Remember that an 'event' was defined as a transition from one state to another state between subsequent titles. Since the distinction between increase and decrease is the most interesting one from a substantive point of view, it was decided to distinguish four types of events in the event history model, namely, *a little increase, a large increase, a little decrease, and a large decrease in the amount of attention between subsequent titles*. We talk about 'little' when we observe only one step: an author changes from one category to the next or to the previous one. When an author moves more than one category forward or backward we refer to it as a 'large' change. This amounts to combining values adjacent to the diagonal elements of table I into the category little changes. Values further away from the diagonal elements refer to large changes.

The fact that the dependent variable was collapsed into a limited number of categories to make an event history analysis feasible can be seen as a disadvantage of the approach that is chosen here. It may seem to lead to some loss of information, but modeling transitions among a small number of categories instead of, for instance, differences in the number of reviews between subsequent points in time, also has advantages. It makes it possible to *define relevant states* and therefore relevant transitions, while

when modeling differences, it is assumed that a difference between, say, 1 and 5 reviews has the same meaning as, say 11 and 15 reviews. Here we only use information on the ordering of different categories of attention. So, in fact, the number of reviews is treated as an ordinal level indicator rather than an interval level indicator.

An event history model for analyzing the debutant authors data has more advantages compared to ordinary regression methods. First, within the event history analysis framework, the fact that we have a different number of titles for each author does not complicate the analysis. We have the possibility of dealing with observations of different length. Second, it is straightforward to model the time dependency of the process and to include time-varying covariates in an event history model. This is important in our case because the variables of which we want to know the effects are all time-varying, that is, their value changes over time. Finally, the event history analysis framework permits the explanatory variables to have different effects on the different types of transitions, which provides more flexibility than performing an ordinary regression analysis on differences scores.

5. The explanatory variables used in the model

The explanatory variables that we used resulted from the formulation of different hypotheses. Van Rees and Vermunt (1996) found that with each successive title it becomes more difficult to get more attention and easier to get less attention. This result can be interpreted by pointing at it as a 'wear out effect'. After an author has published several titles, literary critics may have said all there is to be said on a body of work. Moreover, they can focus their attention on a different group of authors, the new debutants who often belong to a younger generation. With these considerations in mind, it was assumed that there can be observed a turning-point in the amount of attention literary critics pay to an author. We formulated this assumption in two hypotheses: *after an author has published 5 titles, it becomes more difficult to get more attention and easier to get less attention. And: when an author has a seniority of 10 years, it becomes more difficult to get more attention and easier to get less attention.*

When an author who usually writes poetry publishes a novel, that can be seen as a news-item worth paying attention to by literary critics. We must take notice however, that prose-titles in general receive more critical attention than the other genres. So, on the other hand, when an author who usually writes prose, publishes a collection of poems or essays, that may easily escape the notice of reviewers. According to us, a change of genre can have a two-sided effect. We formulated the following hypothesis: *A change of genre will lead to a change in the amount of attention. That change can be either an increase or a decrease.* Finally, it was hypothesized that: *A positive judgment of the previous title leads to an increase in attention, while a negative judgment leads to a decrease in attention.*

With these hypotheses in mind, the explanatory variables that were used in the analysis are the following: *rankorder of the title, seniority, diversity, and mean judgment of the previous title.* The variables were measured and arranged in different categories: rankorder of the title: (1) 2-4 titles (2) 5-9 titles, and (3) 10 or more titles; seniority: (1) less than 5 years active, (2) active between 5 and 10 years,

(3) active between 10 and 15 years, and (4) 15 years active or more; diversity: (1) no change of genre and (2) change of genre, and finally the mean judgment of the previous title:⁵ (1) ≤ 2.5 , (2) 2.5-5.5 and, (3) > 5.5 .

The discrete-time logit model that was used in the current analysis can be described with the following equation:

$$\log(P^d/P^0) = b^d + b^d_{TITL(i)} + b^d_{SEN(j)} + b^d_{DIV(k)} + b^d_{JUDG(l)} .$$

Here, P^d denotes the probability of experiencing event type d , which may be a little/large increase or a little/large decrease, and P^0 indicates the probability of no change in attention between subsequent titles. Thus remaining in the same position is used as the reference category for the dependent variable in the logit equation. In the equation, the log-odds of the probabilities of a little/large increase or a little/large decrease rather than no change in attention are related to TITL (the rank order of the title), SEN (seniority), DIV (diversity), and JUDG (the mean judgment of the previous title). The indices i , j , k , and l denote the levels of the covariates which are treated as nominal variables. As can be seen from the use of the index for the type of event (d) as superscript of the b parameters, the parameters may differ for the various transitions.

After the correct arrangement of the data, the discrete-time logit model can be estimated with standard programs for log-linear analysis. In our case, we used the LEM program (Vermunt, 1993).

6. Results

As in log-linear analysis, the fit of nested models can be compared by means of the *likelihood-ratio chi-square statistic*, denoted by L^2 . The values of L^2 and the number of parameters for some of the models that were estimated are reported in table II.⁶ The easiest way to gain a clear insight in this figure is by comparing the values of the likelihood-ratio chi-square and the number of parameters of the various models. In fact, the models can be tested against each other using differences in L^2 and the number of parameters.

[INSERT TABLE II ABOUT HERE]

First, a model was estimated that contained only the main effects b^d (null model). This model has a L^2 value of 819.3 using 4 parameters (little and large increase, little and large decrease). Next, a model was estimated that contained all effects (full model). This model has a L^2 value of 751.1 with 44 parameters, which is 68.2 lower than of the null model using 40 additional parameters. Thus, we can gain about 68 chi-square points by turning the null model into a full model (note that the lower the value of L^2 , the better the fit of the model). However, the difference of 40 parameters is rather large.

Next, we tried to reduce the number of parameters of the full model by omitting non-significant parameters and by imposing equality restrictions among covariate levels and among transitions. This resulted in models A and B, in which we used 10 and 5 percent significance levels, respectively, to decide whether a particular restriction deteriorated the model fit. Model A which has a few more parameters than model B, reduces seriously the value of L^2 compared to the null model using only 9 additional parameters ($L^2=48.4$; $df=9$; $p=.000$). It also fits significantly better than model B ($L^2=14.1$; $df=5$; $p=.015$). And finally, it does not fit significantly worse than the full model ($L^2=19.8$; $df=31$; $p=.940$).

[INSERT TABLE III ABOUT HERE]

The parameter estimates for model A are reported in table III. The values of the log-linear parameters (b) can be used to determine the direction of an effect and its relative size. For the interpretation of the covariate effects, it is easier to use the multiplicative effects ($\exp(b)$) which are on the scale of the odds rather than the log odds.

The main effects indicate that little changes are more likely to occur than large changes. This is agreement with the results reported in table I. Let us now have a look at the covariate effects on the transition probabilities. For the rankorder of the title, we found a positive effect on a decrease in attention (irrespective of whether it is a little or large decrease) for level 2 and a negative effect on a large increase of attention. The multiplicative effects indicate that the odds of a decrease rather than staying in the same position is 1.52 higher between the 5th and 9th title than before the 5th or after the 9th title. Furthermore, the odds of a large increase rather than no change is 0.50 higher (or 2 times lower) after the 9th title.

The reported parameters for seniority show that when seniority is 4, there is a higher probability of both a little increase and a little decrease. The multiplicative parameters show that the odds of small changes rather than no change are almost 2 times higher (1.86 and 1.88) for authors who are active for more than 15 years than for other authors. The other parameters for seniority indicate that when an author has a seniority of 3, that is, when an author is active between 10 and 15 years, he has a much lower probability of large changes. The multiplicative effects of 0.39 show that the odds of a large increase or decrease instead no change is almost 3 times lower for level 3 of this covariate than for the other levels.

What do these results tell? We assumed that there would be a turning-point in the amount of attention literary critics pay to an author. With the above results that assumption can be confirmed. When we relate the rankorder of the title and the literary age (seniority) of an author to changes in the amount of critical attention, the analysis shows that literary careers have stable and more instable periods. We saw that between titles 5 and 9 ($titl=2$) attention is most likely to decrease. Obviously, this is the more instable period in a literary career. For authors with 10 or more titles ($titl=3$), there is a lower probability

of a large increase. The fact that for authors who are active between 10 and 15 years there is lower probability of large increases or decreases is in agreement with this last result. Authors who have reached this record of service experience more stability in the critical attention they receive. In other words: large fluctuations are less likely when seniority is high.

The b parameter for diversity are relatively high compared to the other covariate effects. A change of genre increases the probability of large increases or decreases, which is, of course, a very general result. This is probably due to the fact that we formulated the idea of a change of genre very roughly. We did not consider the nature of the change we are looking at.

Finally, a non-negative judgment (neutral and positive) increases the probability of a large increase or, the other way around, a negative judgment decreases the probability of a large increase. This is in agreement with the results of Van Rees and Vermunt (1996).

7. Concluding remarks

In this article, we showed how one may gather data permitting to analyze the process of earning a literary reputation in more detail. We also made clear how these data can be rearranged in such a way that it is possible to submit them to an analysis of the kind we now proposed: that is, a logit model shaped in the form of a discrete time event history model. Our analysis shows that there are effects of certain independent variables on the probability of an increase or a decrease in attention for a title, compared to the previous title.

Although the results confirm the hypotheses we formulated, there remain questions to be answered. We found that the effects of the explanatory variables were often two-sided, attention could either decrease or increase. Further research must be more specific on this point. Under which circumstances is attention likely to increase and when is it likely to decrease? Furthermore, questions remain on the issue of a change of genre. Future analysis must make clear which genres exactly we are referring to. Finally, the analysis showed that literary careers have stable and instable periods. We focused on the amount of titles published, and on the amount of active years as a writer. This results must be related to other factors that influence stability and instability in literary careers, for example the productivity of an author and the status of the publisher of a literary work.

Notes

1. This research project is supported by the Netherlands Organization for Scientific Research (NWO).
2. Recent research in the sociology of literature has shown that the attention of literary critics is focused on a very small percentage of all fiction-titles published yearly (Janssen 1994:52).
3. The Dutch Science organization recently started a Humanities research programme entitled 'The impact of using conceptions of literature on quality assignment in the literary field' (Van Rees and Dorleijn 1994; Dorleijn and van Rees 1995). Several projects address the issue how, from a synchronic or diachronic viewpoint, components of conceptions of literature (its core and its prototypical traits) might be constructed into data permitting to answer precise research questions on the material and symbolic production processes in Classical, English, French, German and Dutch literature.
4. This article is based on an IGEL-lecture, held in Budapest in 1994 by Kees van Rees.
5. The judgement was first measured on a 7-point scale: (1) very negative, (2) fairly negative, (3) negative (4) neutral, (5) fairly positive, (6) positive, and (7) very positive.
6. The reported L^2 values are obtained by comparing the model concerned with the saturated model, that is, with the data. If the sample size is large enough, we can use these values to test the absolute fit of the estimated models. This is, however, not the case in the current analysis. Therefore, we only use differences in L^2 values between nested models to compare their fit.

References

- Allison, P.D., "Discrete-time methods for the analysis of event histories." *Sociological methodology* Ed. Leinhardt. San Francisco, CA: Jossey-Bass, 1982: 61-98.
- Bourdieu, P., "The field of cultural production, or: The economic world reversed." *Poetics* 12. 1983: 311-355.
- Janssen, Susanne, *In het licht van de kritiek. Variaties en patronen in de aandacht van de literatuurkritiek voor auteurs en hun werken. [In light of criticism: Variations and patterns of reviewers' attention for writers and their work.]* Doctoral dissertation: University of Tilburg, 1994.
- Rees, C.J. and G.J. Dorleijn, "Literatuuropvattingen in het literaire veld: Over de integratie van twee benaderingen. [conceptions of literature in the literary field: On the integration of two approaches]. *Spektator* 23, 2. 1994: 183-187.
- Van Rees, Kees and Jeroen Vermunt, 1996. "Event history analysis of author's reputation: Effects of critics' attention on debutants' careers." *Poetics* 23, 1996: 317-333.
- Vermunt, J.K., "LEM: Log-linear and event history analysis with missing data using the LEM algorithm." WORC PAPER 93.09.015/7, Tilburg University: 1993.
- Vermunt, J.K., *Log-linear event history analysis: a general approach with missing data, latent variables, and unobserved heterogeneity.* Tilburg: Tilburg University Press, 1996.
- Yamaguchi, K., "Event history analysis." *Applied Social Research Methods*, Vol. 28. Newbury Park, CA: Sage, 1991.

TABLE I: MEAN TRANSITION PROBABILITIES

orig. state	destination state					outflow-table
	0	1-2	3-4	5-6	>6	
0	0.45	0.25	0.20	0.04	0.06	1.00
1-2	0.18	0.33	0.26	0.13	0.10	1.00
3-4	0.18	0.29	0.22	0.17	0.14	1.00
5-6	0.12	0.20	0.23	0.25	0.19	1.00
>6	0.07	0.16	0.24	0.08	0.44	1.00

TABLE II: TEST RESULTS FOR THE MOST IMPORTANT MODELS

		L^2 statistic	# of parameters
null model (only main effects):	819.3	4	
full model (all effects):	751.1	44	
model A (10% level of significance):	770.9	13	
model B (5% level of significance):	785.0	8	

TABLE III: PARAMETER ESTIMATES FOR MODEL A

			b parameter	standard error	exp(b)
main effects	little increase		-0.52	0.16	0.60
	large increase		-1.54	0.32	0.21
	little decrease		-0.73	0.19	0.48
	large decrease		-1.11	0.22	0.33
rankorder title	titl=2, decrease	0.42	0.25	1.52	
	titl=3, large increase		-0.69	0.42	0.50
seniority	sen=4, little increase		0.62	0.35	1.86
	sen=3, large increase		-0.93	0.39	0.39
	sen=4, little decrease		0.63	0.36	1.88
	sen=3, large decrease		-0.94	0.40	0.39
diversity	div=2, large increase		1.01	0.33	2.75
	div=2, large decrease		1.39	0.31	4.01
judgment	judg>=2, large increase	0.59	0.33	1.80	

legend:

rankorder title (1) 2-4; (2) 5-9; (3) >= 10
seniority (1) < 5 years; (2) 5 -< 10; (3) 10 -<15; (4) >=15
diversity (1) no change of genre; (2) change of genre
mean judgment previous title (1) <=2.5; (2) 2.5-5.5; (3) >=5.5