

De logica van de tekstversie in digitaal geschreven literatuur

Dirk Van Hulle

Abstract – This article examines the impact of born-digital literature on editorial theory and genetic criticism, notably on the notion of the textual ‘version’. Born-digital works challenge the logic of the version. At first sight, it seems as if writings whose genesis is recorded with keystroke logging software force us to rethink genetic research so thoroughly that the notion of the version becomes obsolete. However, the research hypothesis of this article is that, on condition that the size of the textual unit to which the notion of ‘version’ applies is clearly defined, this concept does remain a valuable tool in the study of literary geneses. The case study to investigate this research hypothesis is a born-digital story by the Dutch author Ronald Giphart, written with the help of a writing bot.

In 2009 kondigde Marita Mathijssen het einde van de tekstgenetische editie aan: ‘the physical circumstances in which a work comes into being nowadays have changed so much that one can speak of a new era of scholarly editing, and of a radical shift which might well herald the end of the genetic method of editing’ (234). Maar de tekstgenetische oriëntatie in de teksteditie is springlevend, zoals dit artikel zal proberen aan te tonen.

In het onderzoeksgebied van de Writing Studies bestaat al een lange traditie van onderzoek naar digitaal geschreven teksten. In de literatuurwetenschap en de editiewetenschap is dit onderzoek minder ontwikkeld, hoewel er recent binnen de tekstgenetica (*critique génétique* of *genetic criticism*) zeer interessant onderzoek verricht wordt naar mogelijkheden om *digital forensics* toe te passen (Kirschenbaum 2008; Ries 2018). Doorgaans wordt het verschil in aanpak tussen Writing Studies en Genetic Criticism ingegeven door het feit dat de literatuurwetenschap vaak niet met levende schrijvers kan werken. Ook de editiewetenschap ontwikkelt zich gewoonlijk vanuit het werk naar de meest canonieke, niet meer levende auteurs, zoals Shakespeare of Goethe. En toch zijn al deze disciplines gebaat bij meer toenadering.

Dit artikel probeert een stap te zetten in de richting van een dergelijke toenadering door in te zoomen op het begrip ‘versie’ en te onderzoeken of digitaal geschreven (*born-digital*) literatuur ons dwingt om dit concept radicaal overboord te gooien. De stelling van het artikel is dat het begrip ‘tekstversie’ niet noodzakelijk een achterhaald concept wordt wanneer we te maken hebben met digitaal geschreven literatuur. Het digitale medium maakt het niet irrelevant, maar het dwingt ons wel om de werkdefinities van het concept wat scherper te stellen. Het uitgangspunt is dat het begrip ‘tekstversie’ in de editiewetenschap en de tekstgenetica te weinig precies gehanteerd wordt en dat het nodig is om de grootte van de teksteenheid er altijd bij te specificeren. Pas dan kan het opnieuw zijn relevantie bewijzen in het digitale medium. *Born-digital* literatuur noopt ons er dus toe de bestaande, grotendeels structuralistisch geïnspireerde theorieën over dit concept uit de jaren 1970 te herijken.

Het jaar 1971 is in de Duitse *Editionswissenschaft* een mijlpaal. In dat jaar verscheen de bundel *Texte und Varianten*, uitgegeven door Gunter Martens en Hans Zeller. Hierin stelde Siegfried Scheibe de volgende definitie van het begrip ‘tekstversie’ voor: ‘Textfassungen heißen vollendete oder nicht vollendete Ausführungen eines Werkes, die voneinander abweichen. Sie sind durch Textidentität [...] aufeinander beziehbar und durch Textvarianz voneinander unterscheidbar.’ (Scheibe 1971: 17) Zoals Rüdiger Nutt-Kofoth terecht opmerkt, was deze definitie een belangrijke stap omdat ze geen hiërarchie tussen tekstversies suggereert (Nutt-Kofoth 2000: 168). De keerzijde was dat deze definitie in principe

impliceert dat één enkele variant, zelfs één enkele komma, volstaat om van een nieuwe versie te spreken. Theoretisch is dit een waterdichte definitie, maar praktisch ben je er weinig mee als je vorm moet geven aan een wetenschappelijke teksteditie. Terwijl voor theoretici één zwaluw de lente al maakt, loopt het in de praktijk niet zo'n vaart.

1. Werkdefinities: tekst, werk, versie, document

Waar wel een consensus over bestaat, is dat een 'versie' een immateriële eenheid is: 'Bij 'werk' en 'versie' gaat het om immateriële eenheden die een materiële vorm hebben gekregen. Als men uitgaat van het materiaal, hanteert men andere termen. De editor vindt een versie van een werk in een *documentaire bron*: een materiële eenheid die een versie bevat van een werk dat geëditteerd wordt' (Mathijssen 2003: 40). Ook Peter Shillingsburg definieert het document als een materiële eenheid. Van de vier basistermen 'werk', 'tekst', 'versie' en 'document' is volgens Peter Shillingsburg enkel 'document' een concreet object. De overige drie termen zijn abstracte concepten: een 'tekst' is een opeenvolging van tekens; als er meer dan één van deze teksten bestaat, kun je spreken van 'versies'; en het 'werk' is niet de som van al deze versies, maar wordt erdoor geïmpliceerd (Shillingsburg 1996: 176). Shillingsburg wordt – samen met James Thorpe, Hans Zeller, Jerome McGann, Donald Reiman, James McLaverty en J.C.C. Mays – gerekend tot een van de voorstanders van wat Jack Stillinger in zijn 'Practical Theory of Versions' een 'theory of textual pluralism' noemt, die stelt dat 'every individual version of a work is a distinct text in its own right, with unique aesthetic character and unique authorial intention' (Stillinger 1994: 121). Zelfs als iemand een exacte kopie van een manuscript maakt, dan nog kan niet beweerd worden dat de kopie en het manuscript één en dezelfde versie zijn, schrijft Stillinger, 'because the paper, handwriting, occasion, and purpose are entirely different from those of the original' – wat min of meer het editietheoretische equivalent is van de Don Quichot van Pierre Menard in het beroemde verhaal van Borges, 'Pierre Menard, autor del Quijote'.

John Bryant benadrukt dat elke tekstversie een 'identiteit' heeft. Terwijl in de Greg-Bowers traditie vaak van tekstcorruptie werd gesproken (met de implicatie dat het de taak van de teksteditie is om die corruptie ongedaan te maken), ziet de *fluid text theory* – net als de tekstgenetica – het als de taak van de editor om zelfs van zogenaamde corrupte tekstversies de identiteit te bewaren: 'Fluid-text editing preserves "bad" texts in order to trace the causes of a work's entire textual identity' (1045).¹ Volgens Bryant is er een ethische dimensie aan deze taak: 'fluid-text editors have an intellectual and ethical obligation to expose textual interventions by preserving the discrete identities of multiple versions' (1044).² Los van de zeer waardevolle inzichten over de ethische dimensie van deze taak, blijft het echter vanuit editietechnisch standpunt vaag wat nu precies met deze 'versions' of "'bad" texts' wordt bedoeld.

¹ 'The traditional Greg-Bowers approach to editing typically conflates selected fossils into an "eclectic reading text" for critical interpretation, and it shelves the fossil evidence of textual evolution in an apparatus that dis-integrates versions, devalues their textual identity, and obscures the politics of their revision' (Bryant 2010: 1044).

² Vanuit dit ethische standpunt ziet hij 'the problem of mixing textual identities' als vergelijkbaar met andere sociale fenomenen: 'Just as the awareness of the mixings of class, race, ethnicity, gender, and sexual orientation is heightened in any postcolonial, multicultural society, so too must we acknowledge that textual variation reflects dynamic interminglings of power and desire in the marketplace' (1045).

Het begrip ‘versie’ wordt onderscheiden van ‘lagen’³ en ‘fasen’: ‘In een complexe documentaire bron zijn verschillende *lagen* te onderscheiden, omdat de auteur op verschillende tijdstippen met het werk bezig was. Die *fasen* zijn te herkennen doordat er een andere schrijfstof gebruikt is, of doordat de ductus anders is’ (Mathijssen 2003: 47; cursivering toegevoegd). In de definitie van Mathijssen komt een nieuwe ‘laag’ overeen met een nieuwe ‘versie’.⁴ Die tweede versie kan op haar beurt aangepast zijn: ‘als de auteur nogmaals de tweede versie bewerkt, wat blijkt uit nieuwe veranderingen of veranderingen van de varianten uit de tweede werkfase, dan is er sprake van een derde versie’ (279). De vraag is echter of die twee schrijflagen overeenkomen met twee tekstversies. Want wat gebeurt er als twee schrijflagen zeer moeilijk van elkaar te onderscheiden zijn? Als bijvoorbeeld een manuscript in zwarte inkt correcties bevat in dezelfde zwarte inkt, en die tekstlaag wordt gevolgd door een extre revisie in dezelfde inkt, waardoor de twee niet te onderscheiden zijn, dan lijkt het niet mogelijk om ze in de praktijk als twee versies te behandelen.

De definitie van Mathijssen beschouwt de ‘versie’ weliswaar als een immateriële eenheid, maar als de schrijfstof of de ductus criteria zijn om schrijflagen van elkaar te onderscheiden, die vervolgens als aparte versies worden beschouwd, lijkt deze definitie toch een duidelijke band tussen de documentaire bron (de materiële vorm) en de (immateriële) tekstversie te impliceren.

2. Digitale documenten

In het *born-digital* tijdperk gaat die logica niet meer op, of toch slechts gedeeltelijk, zoals Matthew Kirschenbaum opmerkt: ‘Today, the conceit of a “primary record” can no longer be assumed to be coterminous with that of a “physical object”.’ (Kirschenbaum 2013, paragraaf 16) Bovendien schreef Mats Dahlström in 2000 al dat digitale documenten niet materieel gedefinieerd worden: ‘Digital documents are immaterial and therefore logically defined, rather than material and therefore physically defined’ (Dahlström 2000, paragraaf 3, ‘Digitalics’). Hier moet wel onmiddellijk aan toegevoegd worden dat een digitaal document informatie is, en die informatie is ergens opgeslagen op een fysieke drager, zoals een SSD of harde schijf, een RAM-geheugenkaart, een CD-ROM of een USB-stick. Als alle fysieke dragers van het digitale document vernietigd worden, verdwijnt daarmee ook de informatie. Er is een verschil tussen een *digitaal document* (bij voorbeeld een scan, een PDF of een XML-bestand) en de *digitale tekst* op dat document: voor een scan bestaat die digitale tekst uit pixels; voor een XML-bestand bestaat die uit een reeks karakters. Het *document* is in een digitale context dus virtueel: het heeft geen eigen fysieke materialisatie, maar het wordt door middel van software op zo’n manier weergegeven dat het lijkt alsof het die wel heeft.

Ondanks die illusie van materialiteit wordt in een digitale context dus ook wel van ‘documenten’ gesproken (zoals in de .doc extensie van een Microsoft Word-document), maar de definitie is niet dezelfde als die van wat Mathijssen een ‘documentaire bron’ noemt. In

³ ‘Een *laag* bestaat uit een op materiële gronden bij elkaar behorende eenheid in een documentaire bron (of soms in meer documentaire bronnen als een auteur tegelijk op verschillende blaadjes papier werkte)’ (Mathijssen 2003: 280).

⁴ ‘Zodra een auteur naderhand wijzigingen in de grondlaag (inclusief de onmiddellijke correcties) van een teksteenheid gaat aanbrengen, ontstaat er een tweede versie. Soms zal deze te herkennen zijn aan een andere schrijfstof. Er is dan een herkenbare *correctielaag*. Maar het komt ook voor dat er geen verschillen in het schrijfmateriaal te ontdekken zijn en er toch duidelijk in verschillende tijdfasen aan de tekst gewerkt is, zoals blijkt uit doorhalingen, vervangingen boven de regel en toevoegingen al dan niet in de marge. De editeur interpreteert dit dan als een nieuwe *werkfase*, die tot een tweede versie heeft geleid’ (279).

sommige gevallen bewaren schrijvers een versie van elke schrijfsessie. Dat is bijvoorbeeld het geval bij de adaptatie van Dostojewski's *Die Brüder Karamasow* (*The Brothers Karamasow*), geregisseerd door Luk Perceval in het Thalia Theater in Hamburg (2013). Edith Cassiers en Thomas Crombez hebben dit adaptatieproces kunnen analyseren doordat ze inzage kregen in de digitale documenten die ze de 'different versions of the playtext' noemen en die onder de medewerkers uitgewisseld werden tijdens het negen maanden durende repetitieproces (Crombez & Cassiers 2017: 17).

Maar dit is een relatief uitzonderlijke situatie. Vaak schrijft een auteur gewoon verder aan een tekst door het document te overschrijven tijdens een volgende schrijfsessie. Er zijn weliswaar functies zoals 'Revision History' in Wikipedia of 'Version history' in Google Docs die verschillende versies bijhouden, door regelmatig een backup te maken. Daarbij rijst de vraag of een versie dan beter gedefinieerd kan worden in de tijd. Dat zou betekenen dat een versie eerder gekoppeld is aan wat Mathijsen een '(schrijf)fase' noemde dan aan de '(schrijf)laag'. Maar wat is in het digitale tijdperk een 'fase' en hoe wordt die afgebakend? Of wordt een versie bepaald door de willekeurig ingestelde interval waarmee een computerprogramma een backup van een 'document' maakt?

De tussentijdse stand van zaken komt dus hier op neer: de theoretisch waterdichte definitie wordt in de praktijk niet toegepast; de materiële link met schrijflagen in het document valt bij *born-digital* literatuur weg; en de temporele band tussen inhoud en document (bijvoorbeeld het afwerken van een teksteenheid zoals een hoofdstuk, of het afronden van een schrijffase) wordt in de praktijk van het digitale schrijven vaak *de facto* vervangen door een automatisch geprogrammeerde tijdsinterval.

Het valt op dat de meeste definities de term 'tekstversie' beschrijven door het te contrasteren met de termen 'tekst' en 'werk', zelden met het begrip 'document'. Dat de beschikbare definities in het online Lexicon for Scholarly Editing het begrip definiëren door het af te zetten tegen deze andere twee abstracte termen, bevestigt het wezenlijk conceptuele karakter van de versie. Het staat dus los van een tastbaar object, wat betekent dat het niet noodzakelijk afgeschreven is en nog steeds nuttig kan zijn in de context van *born-digital* werken. De vraag is: hoe?

3. *Test case*

Om dat te onderzoeken stel ik voor te vertrekken van een concreet experiment dat onderzoekers van de Universiteit Antwerpen en het Meertens Instituut hebben uitgevoerd in samenwerking met Ronald Giphart (Manjavacas et al. 2017).⁵ De Nederlandse auteur schreef het verhaal 'De mens is de robot van de machine'. Daarvoor maakte hij gebruik van 'Asibot', een schrijfbot die hem de mogelijkheid bood om op gelijk welk moment een zin te vervolledigen door automatisch een syntactisch correcte aanvulling van 100 karakters te genereren. Die aanvullingen kon de bot bovendien in verschillende stijlen schrijven, meer bepaald 8 stijlen, gebaseerd op de werken van enkele Nederlandse en Vlaamse schrijvers als Gerard Reve en Kristien Hemmerechts, op de Nederlandse vertalingen van het werk van Isaac Asimov, en tenslotte ook op de reeds gepubliceerde werken van Giphart zelf. Het

⁵ Het experiment was een samenwerking tussen het Meertens Institute (KNAW, Amsterdam) en ACDC (Antwerp Centre for Digital humanities and literary Criticism, University of Antwerp), m.m.v. Benjamin Burtenshaw, Wouter Haverals, Folgert Karsdorp, Mike Kestemont, Enrique Manjavacas, Vincent Neyt en Dirk Van Hulle.

programma omvatte eveneens *keystroke logging* software, waarmee elke aanslag tijdens het schrijfproces werd vastgelegd. Met deze software begon Giphart een eerste zin te schrijven:

De cameraman keek mij afwachtingsvol aan.

Het belang van het ‘incipit’ wordt in zowat alle vormen van literatuuronderzoek erkend. In de tekstgenetica wordt er wel op gewezen dat het begin van een verhaal of een roman zelden samenvalt met het begin van het schrijfproces (Boie & Ferrer 1993). Ook in het geval van Giphart blijkt de eerste zin van de schrijver niet overeen te zullen komen met de openingszin van de uiteindelijk gepubliceerde versie van het verhaal. In deze eerste versie is de schrijver zich klaarblijkelijk nog heel goed bewust van het nieuwe programma dat op zijn computer is geïnstalleerd en hem als een cameraman voortdurend gadeslaat. Giphart weet dat hij virtueel in de gaten wordt gehouden. In zijn eerste schrijfsessie zwaait hij als het ware even naar de camera: ‘De cameraman keek mij afwachtingsvol aan.’

Voor het schrijven van de tweede zin activeerde hij voor het eerst zijn schrijfbot in de stijl van Gerard Reve. De bot vulde de zin verder aan met 100 karakters. De zin eindigt in het midden van een woord:

‘Ik heb geen verslag van het schip gelezen, meneer, maar ik heb het gevoel dat de bevelen die wij ge

Giphart vervolledigde de zin:

‘Ik heb geen verslag van het schip gelezen, meneer, maar ik heb het gevoel dat de bevelen die wij gekregen hebben niet kloppen.’

Voor het schrijven van de derde zin activeerde Giphart zijn eigen stijl door de schrijfbot om een aanvulling van 100 karakters in de Giphart-stijl te vragen. Opnieuw eindigt de voorgestelde aanvulling in het midden van een woord:

De man achter de balie was ontzet en zei dat hij het niet wist. ‘Dat is niet waar,’ zei hij, ‘en da

Na deze eerste probeersels wiste Giphart alles wat hij en zijn bot tot dan toe hadden geschreven. Na een pauze begon hij opnieuw en typte hij de eerste zin van het verhaal:

Een maand na de dood van Susan Calvin kreeg ik een bericht in mijn inbox.

Uiteindelijk zou die zin veranderen in:

Een *half jaar* na de *atomisering van het lichaam* van Susan Calvin *krijg* ik een bericht in mijn *neurocircuit*. [mijn cursivering, DVH]

Susan Calvin is het hoofdpersonage in veel van de kortverhalen van Asimov. Haar dood is het vertrekpunt voor Giphart. Het verhaal waarvan hij op dat moment enkel de eerste zin heeft geschreven, zal zich afspelen in een maatschappij die op het punt staat door de Machines overgenomen te worden. Het hoofdpersonage is hierover een kritisch pamflet aan het schrijven. Hij weet niet goed wat hij ermee moet doen en heeft hierover drie gesprekken: met een bebaarde humanoïde, met zijn ethiekbots en met een vrouw uit de ondergrondse

verzetsbeweging ‘Genootschap voor de Mensheid’. Dan begint hij weer voort te schrijven, maar het schrijfproces wordt buiten zijn wil om overgenomen door de robot. Dat is in grote lijnen het verhaal dat Giphart op het punt staat te schrijven, met de hulp van Asibot.

Naar het einde van het verhaal toe probeert het hoofdpersoon enkele zinnen toe te voegen aan zijn pamflet. Maar zijn schrijfbot wordt zelfstandiger en begint zelf tekst te produceren. Het hoofdpersoon reageert: ‘Ik laat mijn schrijfbot de tekst schrappen, maar direct komen er nieuwe zinnen.’ Giphart laat die nieuwe zinnen consequent door Asibot genereren. Hij gebruikt hiervoor vooral de ‘Robot Ik’ stem, afgewisseld met de stem van Reve. Ook het hoofdpersoon heeft dus een schrijfbot, maar die blijkt veel autonomer dan de schrijfbot van Giphart.

Buiten de verhaalwereld heeft de tot dusver geschreven tekst op die manier een invloed op de ontwikkeling van het verhaal. Dit is een aspect van wat in Writing Studies de ‘text produced so far’ (afgekort TPSF) genoemd wordt (Flower and Hayes 1980).⁶ Bovendien is het niet alleen de TPSF, maar ook de *manier waarop* die totnogtoe geproduceerd is met behulp van een bot (WTPSF bij wijze van spreken, the Way the Text is Produced So Far) die een impact heeft op het verdere verloop van het verhaal. In het verhaal neemt de machine het hier definitief over met de zin ‘Het is nu genoeg!’ Dit is een bijzonder moment, want net hier beslist Giphart om juist géén gebruik meer te maken van Asibot. Terwijl in het verhaal de schrijfbot de controle overneemt, gebeurt juist het omgekeerde bij Giphart. Hij neemt de controle weer over en schrijft bijna robot-loos het slot van het verhaal, waarin de robot de mens op zijn plaats zet: ‘Wij zijn de toekomst. We zijn voor eeuwig. Wen er maar aan.’

Giphart schrijft zijn verhaal aanvankelijk in de verleden tijd. Maar in een passage waarin het hoofdpersoon overweegt wat hij met zijn pamflet moet doen, begint het gebruik van de verleden tijd tegen te wringen. Giphart schrijft bewust of onbewust een zin in de tegenwoordige tijd, waarna de ‘Robot Ik’ er nog een zin in de tegenwoordige tijd aan toevoegt. Giphart beslist dan in de tegenwoordige tijd verder te schrijven. Op het einde van de alinea keert hij terug naar het begin van het verhaal en verandert hij alle werkwoorden in de tegenwoordige tijd.

Om de tekstgenese te kunnen weergeven heeft het onderzoeksteam het schrijfproces geanalyseerd en herleid tot een selectie representatieve *snapshots*, waarin de meest markante aanpassingen van het verhaal centraal staan. Dit gebeurde op basis van een visualisatie waarin om de ca. 20 seconden *screenshots* van de schrijfapplicatie werden gemaakt die de toestand van de tekst op dat moment precies weergaven. Deze *stills* werden vervolgens achter elkaar gekleefd, zodat een filmische weergave van het schrijfproces mogelijk werd. Om vaart in het verhaal te houden, werden minder relevante beelden tussen de belangrijkste screenshots vooruitgespoeld of verwijderd.⁷

Vergeleken met analoge tekstgenesen, die vaak lacunes vertonen omdat verscheidene tekstversies verloren zijn gegaan, is dit soort digitaal tekstgenetisch materiaal veel

⁶ Met betrekking tot analoog materiaal van niet meer levende auteurs kan op het eerste gezicht de indruk ontstaan dat tekstgenetica enkel met ‘text produced’ (‘TP’ als het ware) gewerkt kan worden, als het materiaal beschouwd wordt als één monolithisch geheel. Maar het is wel degelijk mogelijk om in de schrijfsporen een volgorde op microniveau te onderscheiden en op die manier te achterhalen hoe de tot op een bepaald moment geproduceerde text een invloed heeft op de voortgang van het schrijfproces.

⁷ Een korte documentaire over de tekstgenese door Vincent Neyt en Dirk Van Hulle met *voice over* door Wouter Haverals is online toegankelijk: <https://youtu.be/u-8oSI8ngsQ>.

nauwkeuriger. Terwijl in traditionele tekstgenetica meestal een tekort aan informatie de regel is, lijkt het digitale materiaal bijna een teveel aan informatie te bieden. Elke verkeerde aanslag op het toetsenbord wordt geregistreerd, wat op zich een zeer gedetailleerde dataset oplevert, maar mogelijk ook veel ‘ruis’ oplevert in het globale beeld van de genese.

Een kleine taxonomie van schrijfstappen laat op z’n minst een zevental categorieën onderscheiden op basis van de granulariteit van de teksteenheid: de opeenvolging van tekstversies op het niveau van het verhaal als geheel; de chronologische volgorde waarin afzonderlijke secties (hoofdstukken, alinea’s) van één tekstversie zijn geschreven; revisiecampagnes binnen één tekstversie; de schrijfolgorde van zinnen en woorden; de chronologische volgorde van revisies op zins- en woordniveau; de schrijfolgorde van lettertekens en interpunctie; de chronologische volgorde van revisies op het niveau van lettertekens en interpunctie.

Om het revisiepatroon te analyseren biedt *keystroke logging software* data over het allerkleinste niveau in de taxonomie van schrijfstappen. Binnen de discipline van de Writing Studies wordt via *keystroke logging software* het pauzegedrag, het aantal schrijfsessies en de lengte ervan, het percentage schrappingen, het gebruik van externe bronteksten, de gemiddelde schrijfsnelheid en de vlotheid van het schrijven (‘fluency’) gemeten. Ook in de software-wereld worden zeer verfijnde systemen gehanteerd voor versiebeheer, zoals Git (ontwikkeld door Linus Torvalds), dat niet afhankelijk is van een gemeenschappelijke locatie of een centrale server. Verschillende medewerkers kunnen op die manier aan hetzelfde document werken en de kleinste aanpassing kan getraceerd worden. Tegen de achtergrond van een dergelijke weelde aan data op microschaal kan het op het eerste gezicht lijken alsof *born-digital* werken ons ertoe dwingen om het traditionele concept van de ‘tekstversie’ op te geven en te vervangen door een ander model, bijvoorbeeld het ‘film’-model. Maar de vraag is of dit lineaire model over de hele lijn een verbetering is.

Om dat te onderzoeken verdient het geval van de omschakeling naar de tegenwoordige tijd extra aandacht. Als we van die fase in de genese de ‘film’ laten afspelen, kan het menselijke oog de aanpassingen niet volgen omwille van het beperkt aantal regels dat tegelijk op het computerscherm kan verschijnen. Wanneer Giphart dus midden in het verhaal beslist om terug te keren naar het begin en alle werkwoordsvormen in de tegenwoordige tijd te zetten, maakt de cursor grote sprongen doorheen de reeds geschreven tekst. De reden waarom het voor het menselijke oog zo moeilijk is om die ‘film’ te volgen, is dat twee tijdsassen als het ware over elkaar heen worden gelegd: de chronologie van het narratieve verloop en die van het schrijfproces.

De meest pragmatische oplossing voor dit probleem is het scheiden van die twee tijdsassen, door ze als een x- en een y-as voor te stellen – het systeem dat in de tekstgenetica al decennia wordt toegepast. De vraag is dan hoe dit beproefde systeem aangepast kan worden aan een nieuwe situatie in de literatuur: het feit dat ze digitaal geschreven is.

4. De grootte van de teksteenheid

Wat er vooral aangepast moet worden in de editietheorie is het expliciteren van de grootte van de teksteenheid wanneer over een versie gesproken wordt. Dat heeft op zich niets met *born-digital* literatuur te maken, maar de nieuwe situatie dwingt ons om er eindelijk werk van te maken. Wanneer van een ‘tekstversie’ wordt gesproken, gaat men er vaak stilzwijgend van uit dat het gaat om een versie van de tekst van een werk in zijn geheel. Als dat een roman is,

gaat het om een versie van de hele roman, ook al heeft de schrijver misschien meer dan een jaar aan die ene versie gewerkt. Vanuit chronologisch standpunt is dat een enorm verschil met een versie van een kort gedicht, waar mogelijk slechts een dag aan gewerkt is.

Het is daarom nuttig om met kleinere teksteenheden te werken en een onderscheid te maken tussen versies van een roman, van een hoofdstuk, van een alinea, van een zin of zelfs van een woord. In het *Beckett Digital Manuscript Project*, bijvoorbeeld, wordt gewerkt met zinsversies: om het even waar de lezer zich in een manuscript bevindt, op elk moment kan het nummer van de zin en alle versies ervan worden opgevraagd in een synoptische weergave. Wanneer een zin in een vroege manuscriptversie uiteindelijk niet in de gepubliceerde versie terecht is gekomen, dan krijgt deze zin het nummer van de voorgaande zin (bv. <seg n='234'>) die wel de gepubliceerde versie heeft gehaald, gevolgd door een nieuwe nummering (<seg n='234|001'>).⁸

Deze aanpak is niet alleen op het werk van deze auteur toepasbaar. Ook voor het werk van James Joyce is werken op zinsniveau nuttig. In *Finnegans Wake*, bijvoorbeeld, citeert Joyce in een vroege versie de volgende regel van Byron's *Childe Harold's Pilgrimage* (Canto IV, strofe 179):

Roll on, thou deep and darkblue ocean, roll. (Joyce 1978-79, 56: 2)

In de derde netversie maakt hij een aantal wijzigingen en steekt hij op die manier de draak met het metrum van de jambische pentameter, dat de Engelstalige poëzie eeuwen heeft gedomineerd, door de woorden 'typografisch' overdreven te scanderen:

Rollon thoudeep anddark blueo ceanroll! (Joyce 1978-79, 56: 13)

En uiteindelijk vervormt hij de woorden zodanig dat er van de oorspronkelijke oceaan niet veel overblijft, volgens het principe van het doorfluisterspelletje 'Stille Post' dat model stond voor Joyce's visie op de wereldgeschiedenis – het doorvertellen en dus vervormen van datgene wat we zelf niet hebben meegemaakt:

Rolando's deepen darblun Ossian roll (Joyce 1978-79, 56:174; Joyce 1939, 385)

Voor *Finnegans Wake* is het soms zelfs nuttig om te werken met woordversies om de ontwikkeling van Joyce's zogenaamde 'portmanteau'-woorden te bestuderen (Van Hulle 2005).

Als we het systeem van woord- of zinsversies nu toepassen op een digitaal geschreven tekst als die van Giphart, kunnen we de onoverzichtelijke massa afzonderlijke aanslagen (het resultaat van *keystroke logging*) groeperen in teksteenheden zoals bijvoorbeeld zinsversies. Wanneer Giphart dus op een bepaald moment in het schrijfproces terugkeert naar een zin in het begin van het verhaal om alle werkwoordvormen in de tegenwoordige tijd te veranderen, kan dit moment gemarkeerd worden als een nieuwe versie van die zin (zie Fig. 1).

⁸ In de synoptische weergave verschijnt deze zin dan typografisch gemarkeerd om aan te geven dat hij niet in de gepubliceerde tekst terechtgekomen is.

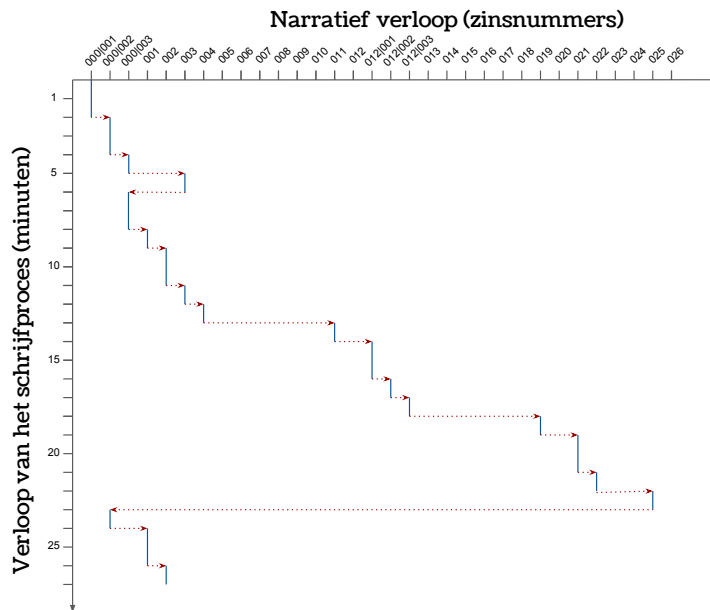


Fig. 1: Visualisering van zinsversies, inclusief zinnen die het niet gehaald hebben in de gepubliceerde versie. De duur van het schrijfproces per zin wordt eveneens gemarkeerd.

Elke zin kan – zoals in het *Beckett Digital Manuscript Project* – genummerd worden aan de hand van een basistekst. Ook de ‘testzinnen’ die Giphart helemaal in het begin schreef (000|001, 000|002, 000|003) kunnen op die manier in kaart gebracht worden. In principe zou elke aanslag op het toetsenbord een nieuwe versie kunnen vormen, maar het gevaar is dan dat de lezer door de bomen het bos niet meer ziet. Mijn voorstel is daarom om een onderscheid te maken tussen een statische versieweergave voor macrogenetisch onderzoek (het bos) en een dynamische weergave voor microgenetisch onderzoek (de bomen).

Een voordeel van *born-digital* werken ten opzichte van analoog geschreven teksten is dat door middel van *keystroke logging* ook de lengte en duur van het schrijven van elke zin gemeten en gevisualiseerd kan worden. Idealiter kan in een editie dus van elke zin een statische weergave (een traditionele transcriptie) en een dynamische (filmische) weergave van alle aanslagen aangeboden worden. De statische weergave maakt macrogenetisch onderzoek (over verschillende versies heen) mogelijk, de dynamische weergave vergemakkelijkt de microanalyse. Dat laatste is bijna niet mogelijk in de reconstructie van analoge schrijfprocessen. Een van de weinige tekstedities waarin dit soort microanalyse bij wijze van experiment is uitgevoerd, is de digitale editie van Willem Elsschots *Achter de Schermen*, waarin voor iedere zin geprobeerd wordt om de opeenvolging van elke afzonderlijke schrijfstap in kaart te brengen. Hierbij komt echter zo veel interpretatie kijken dat het niet mogelijk is om deze praktijk te veralgemenen in de teksteditie.

Op hetzelfde microniveau heheeft ook het begrip *currente calamo* (letterlijk: met lopende veer/pen) herziening. Het gaat daarbij om onmiddellijke varianten, bijvoorbeeld wanneer een auteur die met een ganzeveer, pen of balpen een woord schrijft, het vervolgens doorhaalt en onmiddellijk een alternatief op dezelfde regel schrijft. Als hij deze vervanging net iets later doet (bijvoorbeeld nadat hij het einde van de regel heeft bereikt), is er geen plaats meer op de

regel en zal hij het vervangende woord boven de regel, in de marge of op de tegenoverliggende pagina schrijven. Het heeft weinig zin om bij *born-digital* werken voor dit soort varianten op microniveau nog te blijven werken met een term die nog uit de tijd van de ganzeveer stamt, maar los van de naam is wat het benoemt misschien nog steeds wel relevant. In 2009 klonk Marita Mathijssen erg pessimistisch over de toekomst van tekstgenetische edities. Ze suggereerde dat tekstverwerkingssoftware het begrip ‘Sofortkorrektur’ achterhaald maakte: ‘the whole pre-text has materially vanished. Other, younger writers [...] write and rewrite on the computer and they do not use the memory of the computer as an archive. The work is a whole and the writer is not aware of versions or phases’ (Mathijssen 2009, 237). Maar zoals onderzoekers als Thorsten Ries en Matthew Kirschenbaum hebben aangetoond, maakt digitaal forensisch onderzoek het mogelijk om zeer interessante data voor tekstgenetisch onderzoek en genetische teksteditie terug te vinden. En in de context van digitaal geschreven werken worden onmiddellijke revisies – soms ook ‘point-of-utterance’-revisies genoemd (Lindgren & Sullivan 2006) – nog steeds erkend; met *keystroke logging* is het wel degelijk mogelijk om deze te onderscheiden van andere, latere aanpassingen.

De auteur mag lang en veelvuldig aan zijn zin sleutelen, zolang de wijzigingen plaatsvinden binnen de grenzen van deze zin, kan deze eenheid beschouwd worden als één zinsversie, die statisch weergegeven kan worden in de grafiek. En wie geïnteresseerd is in de interne wijzigingen binnen deze zinsversie, kan dit proces bestuderen in de dynamische, filmische visualisatie ervan. In plaats van een *currente calamo* variant te localiseren wordt het dan mogelijk om te werken met *writing footage* om details van het schrijfproces op zinsniveau te bestuderen. De lengte van deze filmfragmenten wordt in de grafiek weergegeven door middel van de lengte van de verticale streepjes (die elk een zinsversie vertegenwoordigen). Zodra de auteur de zin verlaat om aan een ander deel van de tekst te werken, is die zinsversie afgerond; zodra zij of hij daarna naar de zin terugkeert om er iets in te wijzigen, kunnen we spreken van een volgende zinsversie.

5. Conclusie

Ook voor digitaal geschreven literatuur geldt dus dat één zwaluw enkel theoretisch de lente maakt: als elke aanslag op het toetsenbord aanleiding geeft tot een nieuwe versie, schiet het concept ‘tekstversie’ zijn doel enigszins voorbij. Dit betekent niet dat microgenetische aanpassingen niet geregistreerd hoeven te worden; integendeel, precies omdat we ze in *born-digital* literatuur kunnen registreren en filmisch, dynamisch kunnen weergeven, is het nuttig om de lezer daarnaast ook als houvast een statische weergave aan te bieden. Om het concept ‘tekstversie’ werkbaar te maken in de praktijk, is het nodig om pragmatische oplossingen te zoeken. Net als bij transcriptieconventies staat het elke editeur vrij om een eigen systeem te bedenken. Maar dat veronderstelt wel dat de gekozen conventie geëxpliciteerd wordt. Ook bij het begrip ‘tekstversie’ gaat het om een conventie. De fysieke drager (een bundel papier of een notitieboek) konden vroeger de indruk wekken dat deze materiële sporen de eenheid van een versie bepaalden. In de praktijk komt zo’n overeenkomst tussen materiële drager en tekstversie wel voor, maar ze vallen niet noodzakelijk samen. Waar digitaal geschreven literatuur opnieuw de aandacht op vestigt, is dat tekstversies altijd al een conventie zijn geweest – of in de woorden van John Bryant: ‘versions are critical constructs’ (Bryant 2002: 90).

Het nieuwe medium problematiseert het begrip ‘document’ als iets dat niet langer als een materiële eenheid gedefinieerd kan worden. Dat hoeft geen nadeel te zijn. Integendeel, het vergroot het besef dat document en versie niet noodzakelijk samenvallen. De fijnmazigere

data die keystroke logging software kan bieden, brengen ook meer ‘ruis’ met zich mee en vergroten daarom de noodzaak om zowel binnen de Writing Studies als in tekstgenetica en editiewetenschap te allen tijde te expliciteren wat de grootte is van de teksteenheid waarmee gewerkt wordt. Mits een dergelijke afbakening van de teksteenheid kan het begrip ‘tekstversie’ wel degelijk nuttig ingezet worden voor teksteditie en tekstgenetisch onderzoek van *born-digital* literatuur.

Bibliografie

Beckett 2011 – S. Beckett, *Beckett Digital Manuscript Project*, D. Van Hulle et al. (red.), www.beckettarchive.org, geraadpleegd op 16 juli 2018.

Boie & Ferrer 1993 – B. Boie and D. Ferrer, *Genèses du roman contemporain: Incipit et entrée en écriture*. Paris: CNRS Editions, 1993.

Bryant 2002 – J. Bryant, *The Fluid Text: A Theory of Revision and Editing for Book and Screen*. Ann Arbor: The University of Michigan Press, 2002.

Bryant 2010 – J. Bryant, ‘Rewriting Moby-Dick: Politics, Textual Identity, and the Revision Narrative’. *PMLA* 125.4 (2010), p. 1043-60.

Crombez & Cassiers 2009 – T. Crombez, Thomas & E. Cassiers, ‘Postdramatic methods of adaptation in the age of digital collaborative writing’. In: *Digital Scholarship in the Humanities* 32.1 (April 2017), p. 17–35.

Dahlström 2000 – M. Dahlström, ‘Drowning by Versions’. In: *Human IT* 4.4 (2000), <https://humanit.hb.se/article/view/174/187>, geraadpleegd op 16 juli 2018.

Elsschot 2007 – W. Elsschot, *Achter de Schermen: Digitale Editie*, P. de Bruyn, V. Neyt & D. Van Hulle (red.). CD-ROM. (WEG-serie , 47) ISBN 978-90-72474-69-8. Kalmthout, 2007.

Joyce 1939 – J. Joyce, *Finnegans Wake*. Londen, 1939.

Joyce 1978-79 – J. Joyce, *The James Joyce Archive*, Michael Groden et al. (red.) New York, 1978-79.

Kirschenbaum 2008 – M. Kirschenbaum, *Mechanisms: New Media and the Forensic Imagination*, Cambridge, Mass.: MIT Press, 2008.

Kirschenbaum 2013 – M. Kirschenbaum, ‘The .txtual Condition: Digital Humanities, Born-Digital Archives, and the Future Literary’. In: *Digital Humanities Quarterly* 7.1 (2013), <http://www.digitalhumanities.org/dhq/vol/7/1/000151/000151.html>, geraadpleegd op 16 juli 2018.

Lebrave 2013 – J.-L. Lebrave, ‘Computer forensics: la critique génétique et l’écriture numérique’, *Genesis* 44, p. 137-147

Leijten & Van Waes 2013 – M. Leijten & L. Van Waes, 'Keystroke Logging in Writing Research: Using Inputlog to Analyze Writing Processes', *Written Communication* 30(3), p. 358-392.

Lindgren & Sullivan 2006 – E. Lindgren and K. Sullivan, 'Analysing online revision'. In: K.P.H. Sullivan & E. Lindgren (ed.), *Computer Keystroke Logging and Writing: Methods and Applications* (Oxford: Elsevier), p. 157-188.

Manjavacas et al. 2017 – E. Manjavacas, F. Karsdorp, B. Burtenshaw & M. Kestemont, 'Synthetic Literature: Writing Science Fiction in a Co-Creative Process'. In: *Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)*, collocated with INLG 2017 (Santiago de Compostella, September 4, 2017), p. 29-37

Martens & Zeller 1971 – G. Martens & H. Zeller (ed.), *Texte und Varianten. Probleme ihrer Edition und Interpretation*. München, 1971.

Mathijssen 2003 – M. Mathijssen, *Naar de letter: Handboek editiewetenschap*. 3^e ongewijzigde oplage. Den Haag, 2003. Beschikbaar in de Digitale Bibliotheek voor de Nederlandse Letteren, https://www.dbnl.org/tekst/math004naar01_01/, geraadpleegd op 16 juli 2018.

Mathijssen 2009 – M. Mathijssen, 'Genetic textual editing: the end of an era'. In: *Was ist Textkritik? Zur Geschichte und Relevanz eines Zentralbegriffs der Editionswissenschaft. Beihefte zu Editio* 28. Tübingen: Niemeyer, 2009, p. 233-240.

Nutt-Kofoth 2000 – R. Nutt-Kofoth, 'Schreiben und Lesen: Für eine produktions- und rezeptionsorientierte Präsentation des Werktextes in der Edition'. In: R. Nutt-Kofoth, B. Plachta, H.T.M. van Vliet & H. Zwerschina (red.), *Text und Edition: Positionen und Perspektiven*. Berlin, 2000, p. 165-202.

Ries 2018 – T. Ries, 'The Rationale of the Born-Digital Dossier Génétique: Digital Forensics and the Writing Process: With Examples from the Thomas Kling Archive', *Digital Scholarship in the Humanities* 33(2), p. 391–424.

Scheibe 1971 – S. Scheibe, 'Zu einigen Grundprinzipien einer historisch-kritischen Ausgabe'. In: G. Martens & H. Zeller (red.), *Texte und Varianten. Probleme ihrer Edition und Interpretation*. München, 1971, p. 1-44.

Shillingsburg 1996 – P. L. Shillingsburg, *Scholarly Editing in the Computer Age: Theory and Practice* (3rd Edition). Ann Arbor, 1996.

Stillinger 1994 – J. Stillinger, 'A Practical Theory of Versions'. In: J. Stillinger, *Coleridge and Textual Instability: The Multiple Versions of the Major Poems*. Oxford, 1994, p. 118-141.

Van Hulle 2005 – D. Van Hulle, 'The Inclusion of Paralipomena in Genetic Editions'. In: *Computerphilologie* 5 (2005). <http://www.computerphilologie.lmu.de/jg05/hulle.html>, geraadpleegd op 16 juli 2018.