Readers, Tropes, and Translations: Directions for Digital Research into Youth Literature

Backman, Agnieszka
agnieszka.backman@nordiska.uu.se
Uppsala University, Sweden

Byszuk, Joanna
joanna.byszuk@ijp.pan.pl
Institute of Polish Language of the Polish Academy of Sciences, Poland

Dombrowski, Quinn
qad@stanford.edu
Stanford University, United States of America

Lang, Anouk
anouk.lang@ed.ac.uk
School of Literatures, Languages and Cultures, University of Edinburgh

Murath, Antonia
amurath@zedat.fu-berlin.de
Freie Universität, Department of German and Dutch Philology

Nomura, Nichole
nnomura@stanford.edu
Department of English, Stanford University, United States of America

As Wouter Haverals and Vanessa Joosen note in a recent article on the Dutch children's author Guus Kuijer, there has been little work done at the intersection of digital humanities and children's literature, despite the presence of digitized collections such as the Baldwin Library of Historical Children's Literature, the Auslit digital collection of storytelling, and the Digital Library for Dutch Literature (25). Putting distant reading into the section headed "Unmapped Territories", *The Edinburgh Companion to Children's Literature* (2017) begins the short chapter it devotes to the subject with the tension between the expectation that research into children's literature will situate individual authors within a broad view of the field, and the absence of methodological ways with which to do so (Giddens 305-6).

This panel addresses this gap by bringing together a number of projects which apply computational approaches to the study of youth literature, ranging from translations of popular girls' series, adaptations of the canon, books with well-trodden themes, and medieval texts with young female protagonists. It encompasses a range of approaches to corpus construction, from small-scale studies of the work of a single author with several hundred thousand words, medium-scale projects with corpora spanning multiple languages, to a corpus containing over 30,000 texts. It considers the opportunities made possible by collaboration across different languages, explores what may be learnt by engaging with common metrics from the field of education, considers the challenges of tracing well-known tropes and patterns across versions and reimaginings of texts, and takes up questions of popularity and prestige which bear on the configuration of what Pierre Bourdieu terms the field of cultural production. Bringing together scholars from PhD students to established researchers, and including those who support digital humanities work, the panel surveys how scholars are currently using corpora and computational tools in a subfield of literature in which to date there has not been sustained work in digital literary directions, and to capture the emerging research questions and methodological challenges facing the field.

Stylometric outlook on The Baby-Sitters Club and its translations

Joanna Byszuk

The concept of "the translator's invisibility", i.e. preference of "absence of any linguistic or stylistic peculiarities" (Venuti 1995:1) in translation, has been the cause of much debate, and despite criticism by some translation scholars, it is now considered a core element of translation studies curricula. The concept of authorship in translation is controversial as well - until recently, translators have rarely been recognized on book covers or mentioned in book promotions, and have often struggled to obtain copyright (cf. ibidem: 8-9 and further) or significant profit. Furthermore, on an intellectual level, they have not necessarily been considered creators.

In this talk, we examine the corpus of translations of the Baby-Sitters Club into 11 language versions (distinguishing three French versions, next to e.g. Catalan and Polish) with stylometry, a method of quantitative text analysis based on comparison of texts using differences in word (or other features) frequency distributions, and network analysis. The presentation follows and extends our work on a corpus of 6 language versions (Byszuk & Dombrowski 2022), in which we discovered significant stylometric visibility of ghostwriters, and found that many of the BSC translators have clearly distinguishing styles, although some also had their stylistic impact appear "invisible", evidenced by the same distribution of similarities between texts in the original and translated corpus. The results of this study led us to an observation that translators tend to be more invisible in some language circles, possibly relating to a specific translation tradition, a hypothesis that we further pursue in this paper.

Does paratextual framing lead you right in the translation of the Swedish Baby-Sitters Club books?

Agnieszka Backman

The Swedish Baby-Sitters Club books consist of 17 books, translated by five translators working in three discrete phases over ten years for two publishers. While having many different translators for a series isn't unusual for a children's book series, the change in publishers presents a slightly chaotic history of the translations: from B. Wahlströms förlag for the first six books, to Sabenfeldt/tl*klubben for a reissue of the first book as well as the next five (books 7–11), and then going back to B. Wahlströms for six books from later in the English series' run (books 50, 54, 55, 58, 60, and 63). All three with different series names: Våran klubb, Barnvaktsklubben and The Baby-Sitters Club.

The low status of the translations of girls' books can often be seen in the paratext: out of order translations of series, titles not semantically close to the source text equivalent, and authors' names less faithfully reproduced (Axelsson 2020), of which the first to fit the Swedish Baby-Sitters Club books. Because of their lower status, publishers often allow for bigger changes to the translati-

ons of popular literature and children's literature, e.g., extensive shortening and freer translations. This study uses the Bleualign algorithm for text alignment with English, along with the Swedish spacCy NLP model to investigate if markers of low status literature translation can be seen in the texts of both publishers as it does in the paratexts or if translator practices influence the end product more.

Ambivalence, Ambiguity and Allusion: Getting Beyond 'Bias' in the Chronicles of Narnia

Anouk Lang

Despite impressive advances in text generation, the ability of machine learning to distinguish 'bias'— racist or other forms of prejudiced language—continues to raise concerns. Contending that literary studies has something to teach computer science in this regard, I use the seven-book Chronicles of Narnia (1950-1956) by C.S. Lewis to investigate the way whiteness and nationalism are figured in the texts, and to demonstrate how computational text analysis can identify these aspects. Lewis's widely-read books have considerable resonance with the contemporary moment: as Maria Cecire observes, medievalist children's fantasies such as the Narnia series "helped to establish Britain as a seat of global enchantment and associate 'white' (good) magic with racial whiteness in the minds of young people across the Anglo-American world and beyond" (10). They are also interesting methodologically because, at around 300,000 words, they are too small for machine learning approaches and so must be tackled with less powerful tools, albeit ones more accessible to humanities scholars. What might seem a fairly simple task of searching for patterning around terms like black and white is not as straightforward as it seems, given the ambiguities around whiteness in the books as seen in figures such as the White Witch. Similarly, British nationalism is encoded in the books in ways not easily retrievable by text search. By investigating how these categories can be understood through the metaphorical and figurative relations in the book, I suggest some ways that larger-scale computational approaches might be better equipped to understand and identify bias.

Young Readers, Textual Difficulty, and Genre

Nichole Nomura

In the United States specifically, and Anglophone contexts more broadly, the idea of textual difficulty has been used and abused in a wide range of educational contexts, from underwriting racially-biased tracking to rewarding the kinds of reading educators, parents, and textbook companies designate as valuable and worthwhile. It makes possible quantified comparisons between readers, schools, and districts, and for digital humanists, quantified comparisons between texts.

In this talk, I argue that not only should digital humanists be wary of difficulty metrics like Lexile scores (Smith 1989), the SMOG index (Mc Laughlin 1969), or the Dale-Chall Formula (Dawkins et al 1956), but that digital humanists are perhaps best-positioned to examine what textual difficulty metrics actually capture. This talk uses the task of distinguishing between Children's, Middle-Grade, and Young Adult literature in the Young Readers Database of Literature as a basis for comparison between different difficulty metrics. The various and competing performances of these metrics provides an opportunity to examine their biases and what they actually measure, as well as an opportunity to think about the relationship between assumed audience and assumed skill. Examining the performance of these difficulty metrics across assumed age groups also invites conversations about how other digital humanities methods that depend on similar assumptions about text might or might not perform on literature for young readers. And finally, this task invites critical reflection on the way genres defined by audiences and their assumed abilities can and cannot be defined by proxies for those abilities.

Wherefore Art Thou Shakespeare: How Identifiable are Shakespeare Adaptations for Young Readers?

Quinn Dombrowski

The focus on young people's experiences in many of Shakespeare's plays is one major thread connecting Shakespeare -- arguably the most canonical of the Anglophone literary canon -- to youth literature, a field often siloed from other forms of literary studies. Debates over whether, when, and how youth should read Shakespeare in school have led to an extreme proliferation of adaptations: some simply to facilitate reading, others to make the content more relatable, and yet others to reimagine Shakespearean characters and narratives with a twist on gender, race, or sexuality (Miller 2003). These books are explicit about their Shakespearean origins, which are a non-trivial part of their selling point. But would we be able to recognize them if they did not tell us?

This talk will take a collection of self-identified Shakespeare adaptations for middle-grade and YA audiences -- including adaptations created for students who struggle with reading, modernizations that are intended to retain the original plot and character information, and creative retellings -- as well as a set of texts at similar reading levels with similar topics, and use a variety of text analysis methods (including distance metrics, topic modeling, and stepwise variable selection) to explore the question of how computationally identifiable Shakespearean adaptations actually are, compared to other youth literature that engages with the themes of adolescence. How much of the Shakespearean connection comes out of the use of recognizable character names, compared to a standard group of plot points that can be reduced to a reliable set of vocabulary, or are there other factors at play?

The Handless Queen and Philological Bias: Visualizing a Medieval Euro-Mediterranean Narrative Tradition

Antonia Murath

A girl on the brink of adulthood narrowly escapes her father's sexual advances; she marries a prince, is slandered, re-exiled and ultimately reinstated into a position of power. This is the basic plotline of one of the medieval West's most successful narratives, the *maiden-without-hands*-type. Versions exist across languages, genres and reading communities (noble, mercantile, Christian, Muslim) with surviving manuscripts from the Balkans to North Africa.

Yet, the tradition remains understudied. This might be a corollary of gender and philological bias. The twice-exiled protagonist ultimately triumphs while transitioning from the status of daughter/ child to that of empress/ woman/ saint, without ever fitting the mold of the male knight errant. Meanwhile, the narrative's circulation defies the category of the nation which still demarcates discrete disciplines of literary study; it also defies modern notions of the text as a contained entity. Motif-combinations across versions frustrate efforts to establish a stemma, the most persistent model of textual transmission which indicates a constructed 'original' and its lineage of translated or copied, increasingly 'corrupted' 'derivatives'. Taking seriously Ríkharðsdóttir's proposition to define medieval textuality as "a dynamic corpus […] in constant flux, being refashioned and reshaped" (2012, 8), I propose using multivariate plot analysis to visualize textual relationships differently. My database of more than 120 motifs across 26 versions (Aljemiado, Castilian, Catalan, English, French, German, Italian, Latin) however draws in part on early philologists' data. How, then, does data visualization contribute to generating different insights, opening up new avenues of inquiry, and how might it thus deepen the understanding of the narrative tradition as a whole?

Bibliography

Axelsson, Marcus (2020). Cherry Ames och Vicki Barr kommer till Skandinavien: Utgivningen av två amerikanska flickboksserier ur ett översättningssociologiskt perspektiv. In: *Översättningsvetenskap i praktiken: Om översättningar, översättare och översättande*, eds. Elin Svahn & Lova Meister.

Bourdieu, Pierre. (1993) *The Field of Cultural Production: Essays on Art and Literature*. Translated by Randal Johnson, Columbia University Press.

Bruijn, Elisabeth de, and Mike Kestemont. (2013) "Contrastive Multivariate Analyses of the Middle Low German Flos Unde Blankeflos Tradition." Neuphilologische Mitteilungen, 13, 2: 171–205

Byszuk, Joanna / **Dombrowski, Quinn** . (2022). Stylometric investigations into translationese: The Baby-Sitters Club across languages. In *16th International Conference on Statistical Analysis of Textual Data Proceedings.*

Cecire, Maria Sachiko. (2019) *Re-Enchanted: The Rise of Children's Fantasy Literature in the Twentieth Century*. University of Minnesota Press.

Dawkins, John / **Dale, Edgar** / **Chall, Jeanne S.** (1956). A Reconsideration of the Dale-Chall Formula [with Reply]. *Elementary English* , *33* (8), 515-522.

Edlich-Muth, Miriam. Digitally Mapping Floire et Blanchefleur', blog post https://digitalphilologist.com/digitally-mapping-floire-et-blancheflor/ [04.11.2022]

Giddens, Eugene. (2022) "Distant Reading and Children's Literature", in: *The Edinburgh Companion to Children's Literature*, edited by Clémentine Beauvais and Maria Nikolajeva, Edinburgh University Press, pp. 305–13.

Hagby, Maryvonne. (2017) 'Die Königstochter von Frankreich' des Hans von Bühel: Untersuchung und Edition nach dem Straßburger Druck von 1500. Waxmann.

Haverals, Wouter, and Vanessa Joosen. (2021) "Constructing Age in Children's Literature: A Digital Approach to Guus Kuijer's Oeuvre", in: The Lion and the Unicorn 45, 1: 25–45.

Kiening, Christian. (2000) "Genealogie-Mirakel: Erzählungen vom 'Mädchen ohne Hände'. Mit Edition zweier deutscher Prosafassungen." Geistliches in weltlicher und Weltliches in geistlicher Literatur des Mittelalters, ed. by Christoph Huber et al., De Gruyter, pp. 237–72.

Kinoshita, Sharon. (2014) „Chapter Twenty: Mediterranean Literature." A Companion to Mediterranean History, ed. by Peregrine Horden and Sharon Kinoshita, Wiley Blackwell, pp. 314–329.

Kunkel, Nico. (2023) Modulares Erzählen: Serialität und Mouvance in der Erzähltradition Der „Sieben Weisen Meister". De Gruyter.

Mc Laughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of reading*, *12*(8), 639-646.

Miller, Naomi, ed. (2003) *Reimagining Shakespeare for Children and Young Adults*. New York: Routledge,

Ott, Michael R. (2022) Die Germanistik und ihre Mittelalter: Textwissenschaftliche Interventionen. De Gruyter.

Popović, Pavle. (1908) "Die Manekine in Der Südslavischen Literatur." Zeitschrift für Romanische Philologie 12, 3:312–22.

Ríkharðsdóttir, Sif. (2012) Medieval Translations and Cultural Discourse: The Movement of Texts in England, France and Scandinavia. Boydell & Brewer.

Smith, Dean R. (1989). The Lexile Scale in Theory and Practice. Final Report.

Stein, Robert M. (2009) „Multilingualism" Middle English, ed. by Paul Strohm, Oxford Univ. pp. 23–37

Venuti, Lawrence. (2008). *The Translator's Invisibility: A History of Translation* (2nd ed.). Abingdon, Oxon, U.K.: Routledge.