

Corpus linguistics and disability studies: Ableism, agency, power, and blindness

Sheri Wells-Jensen (Bowling Green State University, US), Jason Wells-Jensen (Bowling Green State University, US), Anastasiia Kryzhanivska (Bowling Green State University, US), Abberley Sorg (Bowling Green State University, US), Kean Kaufmann (independent scholar) and Tony Davis (independent scholar)

One of the purposes of disability studies, in addition to advocating for the rights of disabled individuals, is to provide a thoughtful perspective on works of art and literature with the goal of focusing attention on their portrayals of disabled people, which can both reflect and shape societal attitudes. Attending to these portrayals, then, becomes a means of examining—and in some cases opposing—such attitudes. (See Longmore, 2003, for a useful introduction to disability studies.)

Often, such examinations are focused on agency; i.e., do the disabled characters wield control over their own lives, or are they controlled by others?

A traditional analysis of this sort can be done by selecting scenes or passages from the text and using the tools of literary criticism to explicate them. This approach can obviously lead to controversy; analysts can be accused of ‘cherry picking’ particularly egregious scenes or avoiding portions of the text that appear to contradict their analysis. Here, we employ techniques drawn from corpus linguistics and ranging over an entire text, to establish a firmer basis for claims about a character’s agency.

The novel *All the Light We Cannot See*, by Anthony Doerr (2014), is a particularly cogent example. It has been widely praised, winning the Pulitzer Prize in 2015; in particular, the presence of a blind character, Marie-Laure, has been viewed positively (Faerber, 2015). At the same time, negative reviews of the book cite it as a classic example of destructive stereotypes about blindness (Wells-Jensen, 2016).

Using traditional methods, advocates for the novel could point out passages where the blind main character volunteers for a task to assist the French Resistance and to a scene at the end of the novel where she is portrayed as receiving a graduate degree. On the other hand, critics who judge the book as ableist might point out a quasi-incestuous scene where the father washes Marie-Laure’s hair and numerous occasions where other characters assist her in putting on her shoes or “button her into her coat”. Both sides eventually may come to a standstill as isolated pieces of ‘evidence’ are offered, disputed, reanalyzed, and eventually rejected.

Usefully, constructs such as ‘agency’ have lexical, semantic, and syntactic correlates which can be examined throughout the entirety of the book, avoiding accusations of partisan data mining. These correlates can be located, counted, and discussed with some measure of objectivity. This paper lays out a method for engaging a text in this way.

Thus, the question asked is not “Does this book, as represented by these selected scenes, promote destructive stereotypes of disabled people?”, but rather “Do the words, syntactic frames, and illocutionary forces used consistently

throughout the book to portray this character differ in measurable ways from those representing non-disabled characters appearing in the same text?"

In addition to some simple in-house Perl scripts for sorting and counting words, we used the Stanford CoreNLP 3.7.0 toolkit (Manning et al., 2014) as a starting point. CoreNLP produces parses and pronominal coreference chains, crucial for determining who is involved in events and in what capacity. To assist in the determination of lexical items to be examined, we elected to use the Harvard General Inquirer (General Inquirer, n.d.), as it contains a plethora of categories relevant to the task, including rankings of words as 'strong', 'weak', 'passive', and 'active'. At this point, many detailed investigations became possible:

- We collected conversational turns, associated each with the character responsible for the utterance, and labeled each with its simple illocutionary force. This enabled us to determine, e.g., which characters issued commands or asked for information. Roughly 60% of Marie-Laure's utterances are requests for information, as compared to an average of 28% for other characters. Conversely, only 5% of her utterances are commands, while utterances for other characters range from 10% to 20% commands.
- We also labeled characters' utterances by addressee. Although it is very common for Marie-Laure to ask questions, only 10% of requests for information in the book are addressed to her. Furthermore, half of these are requests for information about her state of health, while the questions addressed to other characters focus on information they possess, events they have witnessed, or the degree to which they agree or disagree with the topic under discussion.
- Using Inquirer words categorized as 'strong' and 'active' which occur frequently in the novel, it is possible to locate their subjects, showing the degree to which the blind character is disproportionately the subject of verbs such as *to feel* and *to beg*. She does not appear as the subject for 'stronger' verbs such as *to explain* or *to attack*, and although there are 18 tokens of the verb *to fight* in the novel, she is the subject of only one of these—its object being *temptation*.

In terms of several linguistic and interactional variables, then, our analysis indicates that the blind character, Marie-Laure, is lacking in agency as compared to other characters in the same novel.

Of course, other variables are at play here; Marie-Laure's age, her gender, her relatively high socio-economic status, and the fact that she is an only child and apparently without friends, all are aspects of her character which co-occur with her blindness. While acknowledging these things, one cannot then ignore the fact that her lack of agency reflects a persistent negative stereotype of blindness. It is remarkable how skillfully and consistently Doerr has—either consciously or unconsciously—crafted the character, forming a reliable lexical and syntactic foundation which supports this effect. If the rendering was consciously done, it is a remarkable example of attention to low level detail. If it was unconsciously done, and the intent was to create a strong character who defies conventional stereotypes of disability, the linguistic pattern of passivity is perhaps even more remarkable.

It is not our intention to create a new kind of imprimatur, a test by which a piece of writing can be either approved as 'inclusive' or dismissed as 'ableist'. What we offer here is a set of reproducible, objective observations and techniques which can be used by both readers and authors to examine the substrate of any piece of writing. What they then choose to do with that knowledge is outside the scope of this paper.

References

- Doerr, A. (2014). *All the Light We Cannot See*. Scribner.
- Faerber, M. (2015). [Review of the book *All the Light We Cannot See*]. *Future Reflections*, 34:1 (Winter 2015). American Action Fund for Blind Children and Adults in partnership with the National Organization of Parents of Blind Children. Retrieved from <https://nfb.org/images/nfb/publications/fr/fr34/1/fr340115.htm>.
- General Inquirer (n.d.). Retrieved from <http://www.wjh.harvard.edu/~inquirer/>.
- Longmore, P. K. (2003). *Why I Burned My Book and Other Essays on Disability*. Temple University Press.
- Manning, C.D., M. Surdeanu, J. Bauer, J. Finkel, S.J. Bethard & D. McClosky. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 55-60.
- Wells-Jensen, Sheri. (2016). On Blindness and the Portrayal of Marie-Laure. In *All the Light We Cannot See*. *LightHouse Interpoint* and *The Toast*. Retrieved from <http://thetoast.net/2016/03/23/on-blindness-and-the-portrayal-of-marie-laure-in-all-the-lightwe-cannot-see/>.