Metaphors we read by: Finding metaphorical conceptualizations of reading in web 2.0 book reviews

Introduction

While interdisciplinary research on metaphor is abundant (Eggs, 2000; Semino & Demjén, 2017; Veale et al., 2016), it is still scarce in Digital Humanities. At the intersection of literary studies, corpus stylistics, and digital humanities, we present an exploratory quantitative metaphor analysis of a corpus of German language lay book reviews. Using a deliberately simple methodological approach that operates on seed words for conceptual sources and targets we investigate how reading experiences of literary texts are metaphorically presented by reviewers. We explore a corpus of approx. 1.3 mill. book reviews for metaphors used to conceptualize the target domain READING EXPERIENCE.

In line with conceptual metaphor theory, metaphors in language are understood as closely linked to human thought processes and experiences (Lakoff & Johnson, 1980, pp. 4–6; Shutova, 2017). They are mappings from typically more basic experiential source domains (LIFE) to more abstract target domains (READING EXPERIENCE), indicated by indirectly used lexis (the words *come, end,* and *road* in "we've come to the end of our road", VUAMC, Steen et al., 2010).

Starting from findings on literature reviews in English (Stockwell, 2009; Nuttall & Harrison, 2018) and on reviews in German (Köhler, 1999), we analyze metaphor patterns in social reading networks, with a particular focus on the mapping READING EXPERIENCE IS MOTION. The main aim at this stage is to draw up a first typology of mappings.

Method and Data

Metaphor Identification

In view of the challenges of reliable automatic metaphor detection (Veale et al., 2016), we apply a deliberately simple rule-based corpus stylistic approach (Deignan & Semino, 2010). A commonly used resource for identification of metaphorical lexical items per source domain is semi-automatic semantic tagging (Demmen et al., 2015). However, in the absence of an out-of-the-box semantic tagger for German, we rely on a 'traditional' onomasiological resource (Dornseiff, 2004). Metaphors are identified by (1) detecting seed words for target domains, (2a-c) detecting source domain seed words in the textual neighborhood of target domain seed words: the metaphor vehicles. Potential metaphors are examined and assigned to a typology of mappings by inspection of KWICs (3).

Step 1. To identify target domain seed words, we compile a list of 'objects of reading experience' (OREs), i.e. noun lemmas that refer to aspects of reading (literary works, such as *Buch* 'book', *Geschichte* 'story', *Roman* 'novel' and parts thereof, such as *Ende* 'ending' or *Spannung* 'suspense', see Table 1).

Step 2. (2a) Potential source domains are pre-identified by manual MIPVU annotation of small samples of the data (cf. Herrmann et al., in press), and the literature on 'reading' metaphors (e.g. Nuttall & Harrison, 2018). For the present paper, we focus on conceptualizations of reading experiences as MOTION (see Herrmann & Messerli, submitted, for metaphor vehicles from the domain FOOD INTAKE). *(2b)* The lexical access points to the MOTION domain are provided by a word list extracted from Dornseiff (2004) for the semantic field *Fortbewegung* (8.3, see Table 2). *(2c)* To find potential metaphor vehicles that refer to ORE (and not to some other referent), cooccurrences are computed between

Table 1: Objects of Reading Experience (ORE) in LoBo

Lemma	Tokens
Buch ('book')	3040,811
Geschichte ('story')	1676,990
Ende ('end/ending')	619,137
Seite ('page/side')	589,534
Charakter ('character')	571,111
Leser ('reader')	559,430
Band ('volume/band')	552,635
Roman ('novel')	512,693
Teil ('part')	505,592
Autorin ('author-F')	502,238
Schreibstil ('writing style')	417,202
Handlung ('plot')	344,847
Spannung ('suspense')	304,937
Autor ('author-M')	280,446
Reihe ('series/row')	267,906
Anfang ('beginning')	266,810
Protagonist ('protagonist')	230,778
Story ('story')	215,364
Thema ('theme/topic')	212,623
Person ('person')	209,701
Kapitel ('chapter')	194,654
Figur ('figure/character')	184,541
Total No. ORE	12,259,980

Table 2: Ten most frequent potential MOTION seed words in LoBo

Lemma	No. of Tokens
gehen ('to go/walk')	685,961
ziehen ('to draw pull drag')	124,291
laufen ('to run')	77,822
lebendig ('alive/lively')	51,841
fliegen ('to fly')	45,756
reisen ('to travel')	41,376
stoßen ('to push')	41,358
fahren ('to drive/ride')	38,166
wechseln ('to change')	35,960
fliehen ('to flee')	29,664
...	...
Total MOTION lemmas	1,481,032

Table 3. Cooccurrences. Most frequent MOTION seed words within a window of ten words of an ORE

Lemma (semantic field: Fortbewegung 'motion')	Freq. within ten words of an ORE
gehen ('to go/walk')	163,320
ziehen ('to draw pull drag')	45,021
fliegen ('to fly')	29,040
lebendig ('alive/lively')	25,743
wechseln ('to change')	14,627
Fahrt ('drive/ride')	14,254
laufen ('to run')	12,613
steigen ('to rise/climb')	11,566
Tempo ('speed')	10,992
starten ('to start')	8,322
...	...
Total Cooccurrences	389,689

'motion' lemmas and ORE, with a window of 10 lemmas around ORE (using raw frequencies, see Table 3).[1]

Step 3. From the resulting frequency list of potential 'motion'-metaphor vehicles (n= 389,689) a sub-section of the most common lemmas is examined by means of KWICs to determine whether potential vehicles were indeed used metaphorically. In a qualitative step, we infer usage patterns from the resulting true metaphor positives (Table 4).

Data

The LoBo corpus (extracted from the social reading platform "Lovelybooks") contains approx. 1.3 mill. German language reviews by 54,000 users, amounting to 439,923,000 words (Table 4), spread over 15 genres. Each review features a rating (1–5 stars) that refers to a specific book. The corpus is lemmatized and PoS-tagged with TreeTagger (Schmid, 1994), and encoded in CWB (http://cwb.sourceforge.net/).

Table 4. Overview of word frequencies of ORE and source domain seed words in LoBo

Type of lemma	Word Count
Target seed words (ORE)	12,259,980
Source seed words (MOTION)	1,481,032
Cooccurrences ORE + MOTION seed word (window=10 w.)	389,689
Total words in corpus	439,923,009

Analysis

A first result is a list of those lemmas from the semantic field *Fortbewegung* 'motion' that occur frequently within a window of ten words of ORE. While it does not yet allow for conclusive results regarding metaphor use, this list serves as an intermediary step towards identifying a multitude of MOTION metaphors for subsequent analysis establishing a typology of mapping patterns.

The analysis of KWICs shows that certain manners of motion are particularly frequent. Notable are the motions of walking, flying, and driving/riding, realized with the lemmas *gehen* 'to go', *fliegen* 'to fly', and *Fahrt* 'ride/drive'. Notable is variance of 'speed', with fast motion *(Fahrt, fliegen)*, and slower motion (*gehen)*.

Another important observation is about agency within the metaphorical scenario. Readers position themselves mainly as (a) observers who see how the plot moves along; (b) agents who actively 'walk' and 'fly' through the story (or a book's pages); (c) patients being put in motion by the book; and (d) companions who travel along with an ORE (see Table 5). Findings

---

[1] Our aim here is not the identification of phraseological units by significance tests against chance distribution (Mutual Information, DICE, and log-likelihood). Rather, raw frequencies allow us to define a window of reference for metaphor vehicles.

demonstrate the complexity of reading that cannot be restricted to passive reception or hedonistic consumption (cf. Rebora et al., 2019).

Table 5. Categories of mappings READING IS MOTION in LoBo

1) an ORE is the moving agent	Die <Seiten fliegen> nur so dahin. 'The pages are simply flying.'
	Category 1 consists of conceptualizations of reading in which the reader is not given an explicit role, but can be inferred to be an observer. An ORE is the central element and is described as moving through space.
2) the reader is the moving agent	Geschrieben ist das Buch wunderbar fließend , man <fliegt förmlich durch die Seiten> 'The book has a wonderful flow (lit. is written wonderfully flowingly), you are positively flying through the pages'
	Category 2 positions the reader as someone who moves through space and an ORE as a different element of that MOTION (e.g. a space through which the reader moves)
3) an ORE as agent moves the reader or another ORE as a patient	Der Spannungsbogen , der sich bis zum Ende hält , hat mich geradezu durch das <Buch fliegen> lassen und ich konnte es kaum aus der Hand legen . 'The arc of suspense that spans until the end has made me positively fly through the book and I almost couldn't put it aside.'
	Category 3 positions an ORE as the instigator of movement of another ORE or the reader.
4) an ORE and the reader move together	Der Leser <pilgert quasi in der Handlung> mit… 'The reader sort of goes on a pilgrimage with the plot…'
	Category 4 conceptualizes reading as a joint motion between reader and ORE.

In all, our study offers a first typology of metaphorical MOTION-mappings in digital shared reading, as well as evidence of the productivity of MOTION as a source domain for READING in German lay reviews (cf. Nuttall & Harrison, 2019, for English reviews). Extending this exploratory phase into statistical analysis, we plan variance analysis with factors as reader's evaluation (star ratings) and book genre (e.g., middle brow vs. popular). Methodologically, we plan to improve precision of metaphor detection, e.g. by including semantic information from resources such as GermaNet, but also through active learning. Generally, further examination of metaphors will allow valuable insight into underlying conceptual and value systems in reader reviews.

References

Demmen, J., Semino, E., Demjén, Z., Koller, V., Hardie, A., Rayson, P., & Payne, S. (2015). A computer-assisted study of the use of Violence metaphors for cancer and end of life by patients, family carers and health professionals. *International Journal of Corpus Linguistics*, 20(2), 205–231.

Dornseiff, F. (2004). *Der deutsche Wortschatz nach Sachgruppen, Mit einer lexikographisch-historischen Einführung und einer ausführlichen Bibliographie zur Lexikographie und Onomasiologie* (8th, fully revised version with alphabetical register, reprint 2010).

Eggs, E. (2000). Metapher, in: Gert Ueding (Hg.), *Historisches Wörterbuch der Rhetorik*, Bd. 5, Tübingen, S. 1109-1183.

Genette, G. (1989). *Paratexte: Das Buch vom Beiwerk des Buches*. (D. Honig, transl.). Frankfurt am Main: Campus.

Herrmann, J.B. & Messerli, Th. (submitted). ...hungere schon nach dem nächsten Band. Eine Untersuchung von Metaphern für Leseerfahrungen in Web 2.0 Literaturrezensionen. 7. Jahrestagung «Digital Humanities im deutschsprachigen Raum» (DhD), Paderborn, D.

Herrmann, J. B., Woll, K., & Dorst, A. G. (in press). Linguistic metaphor identification in German, in: S. Nacey, A.G. Dorst, T. Krennmayr, & W. G. Reijnierse (Hg.), MIPVU in Multiple Languages, Amsterdam: John Benjamins.

Köhler, M. (1999). *Wertung in der Literaturkritik: Bewertungskriterien und sprachliche Ausdrucksmöglichkeiten des Bewertens in journalistischen Rezensionen zeitgenössischer Literatur* (PhD Dissertation, Universitätsbibliothek der Universität Würzburg). Retrieved from https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/784

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Nuttall, L., & Harrison, C. (2018). Wolfing down the Twilight series: Metaphors for reading in online reviews, in: H. Ringrow & S. Pihlaja (Hg.), *Contemporary Media Stylistics*. New York: Bloomsbury Academic.

Rebora, S., Piroska L., & Kuijpers, M. (2018). Reader Experience Labeling Automatized: Text Similarity Classification of User-Generated Book Reviews. In *EADH2018*. Galway: EADH. https://eadh2018.exordo.com/programme/presentation/90.

Schmid, H. (1994). *Probabilistic Part-of-Speech Tagging Using Decision Trees*.

Semino, E., & Demjén, Z. (2017). *The Routledge Handbook of Metaphor and Language*. London, New York: Routledge.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010). *VU Amsterdam Metaphor Corpus*. Retrieved from http://www.ota.ox.ac.uk/headers/2541.xml

Stockwell, P. (2009). *Texture: a Cognitive Aesthetics of Reading*. Edinburgh: Edinburgh University Press.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.