

# ## DIGITAL CLASSICS ONLINE ##

## Linguistik als antifragile Disziplin? Optionen in der digitalen Transformation

Noah Bubenhofer, Philipp Dreesen

**Abstract:** Digitalization has led to massive changes in the sciences, especially in the humanities, as shown by the Digital Humanities. Using linguistics, we discuss these changes on theoretical, epistemological and methodological levels, which consist in particular in the fact that methods of machine language processing are increasingly alinguistic. These are methods that are not based on linguistic theories and do not follow linguistic modelling of language or language usage. It remains to ask how a scientific discipline such as linguistics must be structured so that it becomes antifragile rather than robust in the face of such revolutions. In other words, how it can get richer out of unforeseeable changes and does not primarily have to endure fights of demarcation and defence against such transformations.

### 1. Einleitung

Die Diskussion darüber, was *derzeit* im Zuge der „Digital Humanities“ mit den Geistes- und Kulturwissenschaften passiert, ist ein wichtiger erster Schritt in der Auseinandersetzung mit Digitalisierung in den Wissenschaften und in der Gesellschaft. Nicht weniger relevant scheint angesichts der Risiken und Möglichkeiten im rasanten Fortschritt vor allem digitaler Techniken und Praktiken die Frage der Zukunft von Disziplinen unabhängig von beobachtbaren Effekten der Digitalisierung. Konkret meinen wir damit die Frage zu stellen, was für allgemeine Optionen zur fortwährenden Sicherung es gäbe, damit eine Disziplin und der durch sie konstruierte und bearbeitete Untersuchungsgegenstand weiter Bestand hat. Diese Frage stellen wir vor dem Hintergrund von seltenen, unerwarteten Ereignissen, den sog. ‚black swans‘,<sup>1</sup> wie sie sich z. B. in der digitalen Revolution darstellen. Wir formulieren erste Ideen zu dieser Frage, die lediglich eine Antwortrichtung andeuten, anhand unseres Fachs, der Linguistik. Einschränkend sprechen wir nicht für den stark naturwissenschaftlich ausgerichteten Teil des Faches, sondern denjenigen, den wir in unserem disziplinären (Selbst-)Verständnis wie folgt fassen: Eine theoretisch und methodisch offene Linguistik, die sich für die Zusammenhänge von Sprache und Gesellschaft, Kultur und Medialität interessiert. Hiervon grenzen wir die (eher anwendungsbezogene) Forschungen maschineller Textverarbeitung ab, deren primäres Erkenntnisziel sich stark von unserem unterscheidet, weswegen wir im Folgenden maschinelle Textverarbeitungsforschung als *alinguistisch* bezeichnen. Diese zeichnet sich u.a. dadurch aus, dass sie Digitalität eher in einem ahistorischen und insgesamt weniger reflektierten Sinne verstehen (2).

<sup>1</sup> Vgl. Taleb, Nassim Nicholas: The black swan: the impact of the highly improbable, New York: Random House 2007.

Um diese Tendenzen einordnen und bewerten zu können, ist es erforderlich, von einer adäquaten Beschreibung von Wissenschaft und Disziplin im Zuge von DH auszugehen (3). Diese Schritte bereiten die Erörterung der Frage vor, was eine Disziplin antifragil macht, d.h. welche Optionen die Linguistik bereithalten muss, um aus unbekanntem Ereignissen und Entwicklungen in der Zukunft gestärkt hervorzugehen (4). Dies wird im Fazit noch einmal pointiert (5).

## 2. Aktuelle Tendenzen im linguistischen Feld

Aus linguistischer Sicht könnte die Aussicht in die digital durchdrungene Welt nicht besser sein: Die Massen von Daten, die tagtäglich entstehen, bestehen zu großen Teilen aus Text, aus sogenannten «unstrukturierten Daten», wie sie aus informatischer Sicht von den «strukturierten Daten» unterschieden werden. Strukturierte Daten, wie beispielsweise die meisten Metadaten eines Tweets (Datum, Zeit, Autor\_in, Geokoordinaten etc.), können von Maschinen relativ gut verarbeitet werden, da der zu erwartende Datentyp und die möglichen Inhalte bekannt sind und es vergleichsweise einfach ist, daraus beispielsweise ein Interaktionsnetzwerk abzuleiten (wer, wann, mit wem kommuniziert hat). Unstrukturierte Daten wie Text folgen jedoch syntaktischen, semantischen und pragmatischen Regularitäten, die für eine maschinelle Analyse zunächst decodiert werden müssen.

Trotz der Schwierigkeiten maschinellen Umgangs mit solchen Daten ist die Bedeutung von Sprachdaten unbestritten. So etwa im Rahmen von wirtschaftlichen Interessen, um Meinungen und Bewertungen von Konsument\_innen, die in Textform abgegeben werden, maschinell auszuwerten. Dazu werden beispielsweise Methoden der Sentiment-Analyse eingesetzt, mit der versucht wird, bewertende Texte wie Produktebewertungen automatisch in die Klassen positiv, negativ und neutral einzuteilen. Auf große Datenmengen hin angewendet, können Rückschlüsse daraus gezogen werden, wie Produkte angenommen und bewertet werden, um die Marketingmaßnahmen entsprechend anzupassen. Andere Beispiele für die Bedeutung der maschinellen Textanalyse sind die inhaltliche Erschließung von riesigen Textdokumenten-Archiven, um die Dokumente etwa automatisch nach Themen zu gruppieren, oder die maschinelle Übersetzung, um Textdaten rasch in beliebigen Sprachen verfügbar zu machen.

Die Bedeutung von Texten wird also auch außerhalb der Linguistik oder der Philologien allgemein erkannt. Der «linguistic turn» ist also nach geraumer Zeit in der Mitte der Gesellschaft angekommen.

Allerdings ist bei den Methoden der maschinellen Textanalyse ein fundamentaler Wandel zu beobachten, der die Rolle der Linguistik berührt. Das lässt sich an der Geschichte der Computerlinguistik gut zeigen, bei der ein eindeutiger Trend von sogenannten regelbasierten Ansätzen zu statistischen Ansätzen und jüngst zu Systemen neuronalen Lernens zu beobachten ist.<sup>2</sup> Bei regelbasierten Ansätzen wird versucht, Texte mittels linguistischer Kategorien zu annotieren und über verschiedene Regeln, die diese linguistischen Annotationen nutzen, maschinelles «Verstehen» zu ermöglichen. Grundlage dafür sind linguistische Theorien zu Grammatik, Semantik, Pragmatik etc. Bei der Aufgabe, bewertende Texte nach positiv-neutral-negativ zu

---

<sup>2</sup> Uszkoreit, Hans: „Linguistics in Computational Linguistics: Observations and Predictions“ Athens, Greece: Association for Computational Linguistics, 03.2009, <http://www.aclweb.org/anthology/W09-0105>.

klassifizieren, würde man beispielsweise mit nach einer bestimmten linguistischen Theorie vorgenommenen Wortlisten arbeiten, eventuell in Kombination mit Regeln, die die syntaktische Struktur betrachten (etwa um Negation zu berücksichtigen).

Der statistische Weg geht von manuell annotierten oder klassifizierten Daten aus, deren Eigenschaften «maschinell gelernt» werden. Dabei arbeitet man mit einem Set von Merkmalen, von denen man annimmt, dass sie helfen, die Daten zu klassifizieren. Das statistische Lernen besteht dann darin, die Merkmalsausprägungen zu finden, die die Daten und ihre Klassifikation am besten beschreiben. Ein solches statistisches Modell kann dann auf nicht-klassifizierte Daten angewandt werden, um sie maschinell zu klassifizieren. So dienen etwa manuell kategorisierte Daten von positiven und negativen Produktbewertungen dazu, die statistisch auffälligen Merkmalsausprägungen zu lernen, etwa ein bestimmtes Vokabular, das typisch für positive oder negative Bewertungen ist. Linguistische Theorien spielen bei der Auswahl der Merkmale eine Rolle, deren Ausprägungen fürs Lernen verwendet werden.

Eine seit etwa zehn Jahren besonders wichtig gewordene Methode ist das maschinelle Lernen mittels neuronaler Netze, sogenanntes «Deep Learning». Der Unterschied zu den klassischen Verfahren maschinellen Lernens über mitunter komplexe und zahlreiche Merkmale liegt darin, dass oft nur sehr basale Merkmale berücksichtigt werden, bei Texten etwa Buchstaben-n-Gramme, und dass das Lernen in einem neuronalen Netz geschieht. Abstraktere Merkmalsausprägungen sollen sich dabei, so die Erwartung, als Muster aus den simplen Merkmalen ergeben. Das Lernen findet zudem auf mehreren verborgenen Ebenen statt, so dass das Lernen nicht direkt beobachtet und auch die Frage, welche Merkmale nun welchen Einfluss auf das gelernte Modell haben, kaum beantwortet werden kann.<sup>3</sup> Deep Learning ist für viele Anwendungen sehr erfolgreich, was beispielsweise die großen Fortschritte in der maschinellen Übersetzung zeigen, etwa beim kommerziellen Angebot «DeepL» ([www.deepl.com](http://www.deepl.com)).<sup>4</sup>

Entscheidend für unsere Diskussion ist nun aber, dass solche Methoden, obwohl in vielen Fällen sehr erfolgreich, gar nicht auf linguistische Theorien zurückgreifen müssen, also alinguistisch vorgehen. Linguistinnen und Linguisten werden dafür also nicht gebraucht. Übertragen auf die Geistes- und Kulturwissenschaften bedeutete dies, dass geisteswissenschaftliche Theorien und Methoden an Bedeutung verlieren, wenn es um die Analyse und Interpretation von kulturellen Artefakten geht.<sup>5</sup>

Die Digital Humanities sind ein offensichtlicher Versuch, diesen Konflikt, wenn nicht zu lösen, zumindest zu diskutieren und mit diesem Konflikt auch zu experimentieren, indem auch alinguistische Methoden wiederum für geistes- und kulturwissenschaftliche Forschungsinteressen genutzt werden. Dabei wird der Einsatz dieser Methoden kritisch reflektiert.<sup>6</sup> So kann etwa problematisiert werden, dass die unterschiedlichen Forschungsinteressen von alinguistischen Disziplinen und der Linguistik zu unterschiedlichen Bewertungen der Methoden führen: Aus Sicht eines Unternehmens, das positive von negativen Produktebewertungen scheiden will, ist

---

3 Goodfellow, Ian, Yoshua Bengio und Aaron Courville: *Deep Learning*, MIT Press 2016, S. 5–6, <http://www.deeplearning-book.org>.

4 Konsequenterweise ließen wir das Abstract für den vorliegenden Beitrag von deepL übersetzen, wobei wir den Ausdruck «upheaval» für «Umwälzung» an zwei Stellen durch die ebenfalls von deepL vorgeschlagenen Übersetzungsvarianten «revolution» und «transformation» ersetzten.

5 Bubenhofer, Noah: „Wenn ‚Linguistik‘ in ‚Korpuslinguistik‘ bedeutungslos wird. Vier Thesen zur Zukunft der Korpuslinguistik“, *OBST. Osnabrücker Beiträge zur Sprachtheorie* 1 (2018).

6 Ebd.

eine Methode dann gut, wenn sie diese Aufgabe gut löst – gemessen an einem sogenannten Goldstandard, also den manuell klassifizierten Testdaten.

Aus linguistischer Perspektive ist eine Methode aber dann «gut», wenn sie dazu beiträgt, eine Funktion wie «Bewerten» besser zu verstehen. Ein Verfahren maschinellen Lernens mit theoriegeleitet ausgewählten Merkmalen und einem transparenten statistischen Modell mag dieses Ziel erfüllen, weil nachvollzogen werden kann, welche Merkmalskombinationen bestimmte Texteingenschaften besonders gut charakterisieren. Schwieriger ist dies bei intransparenten «Black Box»-Systemen, etwa beim Deep Learning, da ja gerade nicht eingesehen werden kann, welche Merkmale für die Charakterisierung der Texte entscheidend waren.<sup>7</sup>

Zusammenfassend kann also gesagt werden, dass Digitalität, aufgefasst als Ensemble von Computerisierung, Verdichtung, Praktiken des Umgangs mit digitalen Artefakten, digitalen Medien, Code, Algorithmen und Effekten der Skalierung, einen enormen Einfluss nicht nur auf Wirtschaft und Gesellschaft, sondern auch auf die Wissenschaft hat. Für die Philologien ist dabei insbesondere der Effekt von Digitalität auf die Produktion, Rezeption und Analyse von sprachlichen Daten von Interesse. Man könnte Digitalität also als Black-Swan-Ereignis auffassen, dessen Auftreten und dessen Auswirkungen nicht vorhersehbar gewesen sind, obwohl Digitalität und die Art digitaler Transformationen wie Sortieren, Selektieren und Rekombinieren als grundsätzliche Kulturtechniken eine lange Tradition haben (vgl. die Thesen von Sybille Krämer in diesem Heft). Mit dem Computer, aufgefasst nach Kay und Goldberg<sup>8</sup> als Metamedium, das fast alle anderen Medien sein kann, und der enormen Geschwindigkeit, mit der digitale Transformationen durch den Einsatz von Transistoren anstelle von Elektronenröhren und Relais möglich werden,<sup>9</sup> ergaben sich jedoch unvorhersehbare Effekte.

Diese Dimensionen von Digitalität, die bei alinguistischen Ansätzen normalerweise nicht reflektiert werden, könnte von den Geisteswissenschaften produktiv aufgegriffen werden. Um diesen Gedanken zu entwickeln, ist es zunächst erforderlich, Forschung und wissenschaftliche Disziplin in der digitalen Transformation adäquat zu erfassen.

### 3. Disziplinäre Transformationen: Denkstile der Linguistik

Wir möchten die von Sybille Krämer eingeführte (vgl. die Thesen von Sybille Krämer in diesem Heft) und für diese Ausgabe von Digital Classics Online aufgegriffene Metapher „Stachel des Digitalen“ nutzen, um mit ihr Überlegungen zur darin enthaltenen Verletzlichkeit der Linguistik anzustellen. Grundsätzlich transportiert die Metaphorik des ‚Stachels im Fleisch‘ ein recht einseitiges Bild von Transformationsprozessen in der Wissenschaft: Erstens handelt es sich um eine Metapher aus der Perspektive der Leidtragenden; der Stachel quält und drängt diejenigen, die der digitalen Revolution skeptisch gegenüberstehen. Es stellt sich aber die Frage, ob diese Personen ‚das Digitale‘ so empfinden oder ob sie Entstehungen wie DH lediglich zur Kenntnis nehmen, da sie sich davon nicht betroffen fühlen.

<sup>7</sup> Dies ist eine zugespitzte Aussage. Es gibt beschränkte Möglichkeiten, einem neuronalen Lernsystem zuzusehen, wie es lernt und daraus Schlüsse zu ziehen. Olah, Chris u. a.: „The Building Blocks of Interpretability“, Distill (2018). Bei Bilddaten funktioniert dies gut (vgl. <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>, 31. Mai 2018).

<sup>8</sup> „Personal Dynamic Media“, *Computer* 10/3 (1977), S. 31–41.

<sup>9</sup> Lévy, Pierre: „Die Erfindung des Computers“, in: Serres, Michel (Hrsg.): *Elemente einer Geschichte der Wissenschaften*, 2. Aufl., Frankfurt am Main: Suhrkamp Verlag 1998, S. 905–944, hier S. 910.

Zweitens ist ‚Stachel im Fleisch‘ stets additiv und statisch gedacht: Es ist nicht die Linguistik als Teil der traditionellen Geisteswissenschaft, die seit einigen Jahren den „Stachel des Digitalen“ spürt – vielmehr wandelt sich die Linguistik mit den digitalen Methoden in Teildisziplinen als Fach, am besten zu sehen und zu belegen an der Korpuslinguistik. Neue Impulse („turns“, Paradigmen, Theorien, Methoden, Techniken, Quellen, soziale Relevanz etc.) bleiben also nicht aussen vor, sondern werden im Fach adaptiert, modifiziert und in einigen Fällen kanonisiert.

Wir schlagen eine andere – nicht weniger metaphorische – Vorstellung von Transformationsprozessen in der Wissenschaft vor. Mit Ludwik Flecks Theorie vom Denkstil (1935/1980) nähern wir uns dem Thema wissenssoziologisch über die Erkenntnispraktiken:<sup>10</sup> Erstens sind Erkenntnisprozesse stets zu historisieren, zweitens „manifestieren [sie] sich in Erkenntnissen, in menschlichen Aktivitäten und Handlungen“<sup>11</sup> und drittens finden [sie] in einem sozialen Umfeld statt, den Fleck als *Denkkollektiv* bezeichnet, d.h. eine „*Gemeinschaft der Menschen, die im Gedankenaustausch oder in gedanklicher Wechselwirkung stehen*“.<sup>12</sup> Die „Denkkollektive“ stellt sich Fleck als relativ und prozesshaft veränderbar vor, sowohl hinsichtlich ihrer Quantität (Menge der Mitglieder) als auch ihrer Qualität (Handlungsformen). Er macht deutlich, dass das Kollektive immer zugleich ein Produkt von Handlung und Kommunikation und einen Prozess des Handelns und Kommunizierens darstellt.<sup>13</sup> Da wir mehreren und oft sehr unterschiedlichen Denkkollektiven angehören, etwa einem vom skeptischen, theorieaffinen Denken geprägtem Kreis<sup>14</sup> und einem entschlossenen Projekt der angewandten Wissenschaft,<sup>15</sup> führt das zu folgendem Mechanismus: „Je größer die Differenz zweier Denkstile, um so geringer der Gedankenverkehr“.<sup>16</sup> Innerhalb der Geisteswissenschaft besteht zwischen DH-affinen und traditionellen Forschungsgruppen wenig Austausch, sowohl personenbezogen wie auch auf Tagungen, in Publikationen und in Form gemeinsamer Forschungsprojekte und Studienangeboten. Es bleibt somit auch die wissenschaftliche Kontroverse aus, wenngleich gerade diese Interaktionsform wichtig wäre, um die Denkstilgebundenheit bewusst zu machen.<sup>17</sup> Den Umstand des geringen Austausches erklärt Fleck mit dem Konzept ‚Denkstil‘, der die geistig-praktische Tätigkeit im Erkenntnisprozess erfasst: Im Denkkollektiv wird wissenschaftliches Handeln zum Zweck der Ergebnissicherung und zur Entdeckung von Neuem koordiniert, sodass das Wahrnehmen von Untersuchungsgegenständen, Erkenntnissen und Fehlschlägen in

---

10 Vgl. zum Folgenden Andersen, Christiane u. a.: „Erkenntnis als soziale Praxis : Ludwik Flecks Wissenschaftstheorie aus sprachwissenschaftlicher Sicht“, *Denkstile in der deutschen Sprachwissenschaft : Bausteine einer Fachgeschichte aus dem Blickwinkel der Wissenschaftstheorie Ludwik Flecks* (2018), S. 11–65.

11 Ebd., S. 14.

12 Fleck, Ludwik: Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv, hg. von Lothar Schäfer und Thomas Schnelle, Frankfurt am Main: Suhrkamp 1980, S. 54; Herv. i. Orig.

13 Andersen u. a.: „Erkenntnis als soziale Praxis“, S. 18.

14 Vgl. Dreesen, Philipp: „Denkstimmung der Skepsis : ein Beschreibungsansatz zur Entstehung der Diskurslinguistik“, *Denkstile in der deutschen Sprachwissenschaft : Bausteine einer Fachgeschichte aus dem Blickwinkel der Wissenschaftstheorie Ludwik Flecks* (2018), S. 167–190.

15 Vgl. Stücheli-Herlach, Peter, Maureen Ehrensberger-Dow und Philipp Dreesen: „Energiediskurse in der Schweiz: Anwendungsorientierte Forschung eines Kommunikationsfelds mittels digitaler Daten“, *Working Papers in Applied Linguistics* (im Druck).

16 Fleck: Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv, S. 142.

17 Vgl. Andersen u. a.: „Erkenntnis als soziale Praxis“, S. 58–62.

spezifischer Weise gerichtet wird.<sup>18</sup> Der Denkstil verbindet die Praktiken-Dimension wie Geräte (Wahl der Software, Datentyp) und Fachtermini<sup>19</sup> mit der epistemologischen Dimension wie dem erlernten ‚Sinn-Sehen‘ als Prozess des Gestaltwahrnehmens.<sup>20</sup> Das Erlernen bringt mit sich, „daß wir mit dem Erwachen der Bereitschaft, bestimmte Gestalten wahrzunehmen, die Fähigkeit verlieren, anderes wahrzunehmen“.<sup>21</sup> Es ist somit denkstilbedingt, ob wir in einem Stapel Tagebücher den Ausdruck der kulturellen Praktik des Tagebuchschreibens erkennen oder eine unstrukturierte Datensammlung: Je nach dem wird uns dies zu von uns für mehr oder weniger zwingend erachteten Praktiken führen, d.h. die Beschreibung des Untersuchungsgegenstands und seine methodische Bearbeitung (z. B. im Sinne der Grounded Theory offen zu kodieren oder korpuslinguistisch mittels Annotation zu strukturieren).

Zusammengefasst: „Denkstile können also als *disziplinär geformte, sozial angeeignete und inkorporierte Forschungspraktiken* verstanden werden, als in bestimmten Kollektiven (den *Denkkollektiven*) akzeptierte Formen der Erkenntnisfindung und Erkenntnisdarstellung.“<sup>22</sup> Mit dem Denkstilansatz kann die digitale Transformation als Frage der Veränderbarkeit prägender Denkstile und Denkkollektive bei fortwährender Existenz des eigenen heterogenen Fachs aufgefasst werden. Dabei kann zunächst konstatiert werden, dass viele linguistische Denkkollektive digitale Methoden lediglich zur Kenntnis nehmen: Da sie dem eigenen Denkstil nicht entsprechen, können sie kaum als erkenntnisfördernde Praktik in Betracht gezogen werden. Da aber der Denkstil im Kollektiv aktiv erworben werden muss, haben vermutlich neue, zum bestehenden Denkkollektiv hinzukommende Mitglieder in der Vermittlung digitaler Denkstile (z. B. Programmier- und Statistikkompetenzen, Zugriff auf Korpora) einen schweren Stand. Auf der anderen Seite gibt es bereits erste Veränderungen in den Denkstilen: Da ein Großteil unseres Sprachgebrauchs in digitaler Form (vor allem im WWW) stattfindet, werden auch die Erkenntnisinteressen, Theorien, Methoden, Geräte und Begriffe daran angepasst.

Die trotz unserer Reflexion von uns mit diesem Text erneut ansatzweise vollzogene Fragmentierung von Denkstilen und Denkkollektiven aufzulösen, ist ein wichtiger Schritt in der Überlegung zu einer Linguistik als antifragile Disziplin.

## 4. Antifragilität und die Zukunft der Linguistik

Nassim Taleb weist mit seiner Überlegung zum Erscheinen von ‚black swans‘ in der Geschichte auf einen wenig berücksichtigten Aspekt unseres planerischen Denkens und Handelns hin: Wir sind davon überzeugt, dass wir verstehen, was derzeit um uns herum passiert; wir überbewerten unsere Informationen und deren Verwertung; und wir erklären historische Ereignisse *ex post* als nicht-kontingent.<sup>23</sup> Unser Bedürfnis, Prognosen zu erstellen, verhindert regelmä-

18 Vgl. Fleck, Ludwik: „Wie entstand die Bordet-Wassermann-Reaktion und wie entsteht eine wissenschaftliche Entdeckung im allgemeinen?“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011, S. 181–210, hier S. 121.

19 Vgl. Fleck, Ludwik: „Das Problem einer Theorie des Erkennens“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011, S. 260–309, hier S. 296f.

20 Vgl. Andersen u. a.: „Erkenntnis als soziale Praxis“, S. 38.

21 Fleck, Ludwik: „Schauen, sehen, wissen“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011, S. 391.

22 Andersen u. a.: „Erkenntnis als soziale Praxis“, S. 34; Herv. i. Orig.

23 Vgl. Taleb: *The black swan*, S. 8.

Big, dass wir unvorhersehbare Dinge in Betracht ziehen. So berücksichtigen wir Schwarze Schwäne nicht ausreichend, d.h. „große, unvorhersehbare, irreguläre Ereignisse“,<sup>24</sup> die zwar selten und höchst unwahrscheinlich sind, die aber bei ihrem Eintreten weit- und tiefgreifende Konsequenzen haben. Mag die Entdeckung des Penizillins, der Aufprall eines Tsunami oder der Zusammenbruch der Sowjetunion zwar als Ereignis unwahrscheinlich bzw. selten gewesen sein, so sind gleichwohl unter anderem die volkswirtschaftlichen Auswirkungen Jahrzehnte spürbar. Das Eingeständnis des eigenen Unvermögens, solche Ereignisse vorzusehen, darf laut Taleb nicht dazu führen, bessere Prognosen erstellen zu wollen. Vielmehr bedeutet es in der Konsequenz, sich mit der eigenen Unwissenheit zu arrangieren: „Black Swan logic makes *what you don't know* far more relevant than what you do know.“<sup>25</sup> Taleb sieht auch das Internet als black swan an, da niemand mit seiner rasanten tiefgreifenden Veränderung in Beruf und Alltag gerechnet hat.<sup>26</sup> Somit kann, so möchten wir Taleb hier weiterdenken, auch die digitale Revolution in den Wissenschaften und mit ihr die Digital Humanities als black swan aufgefasst werden.

Unter anderem zum Management von Schwarzen Schwänen entwickelt Taleb das Konzept und den Ausdruck *antifragile*, um im Risikomanagement die Zukunftsfähigkeit von Institutionen bewerten zu können. Während robuste Institutionen lediglich stabil genug sind, durch eintretende Risiken nicht geschwächt zu werden, profitiert die antifragile Institution vom Risiko. Seine Definition von *antifragile* lautet: „Some things benefit from shocks; they thrive and grow when exposed to volatility, randomness, disorder, and stressors and love adventure, risk, and uncertainty. Yet, in spite of the ubiquity of the phenomenon, there is no word for the exact opposite of fragile. Let us call it antifragile. Antifragility is beyond resilience or robustness. The resilient resists shocks and stays the same; the antifragile gets better.“<sup>27</sup> Zum besseren Verständnis von Antifragilität bietet Taleb die mythologische Figur der Hydra an, der für jeden verlorenen Kopf stets zwei Köpfe nachwachsen; im Gegensatz dazu ist der aus seiner eigenen Asche unverändert aufsteigende Phönix lediglich robust.<sup>28,29</sup> Wichtig ist die ethische Maxime, dass es bei Antifragilität nicht um die Verbesserung des eigenen Zustands auf Kosten der Fragilität anderer geht.<sup>30</sup>

Die Linguistik wird grundsätzlich antifragil, wenn und indem sie mehr Optionen erhält. Anders gesagt: Bedingung für Antifragilität sind weder Überlegungen zum Status der Linguistik hic et nunc noch Prognosen über mögliche anstehende Denkstil- und Paradigmenwechsel. Bedingung für Antifragilität ist auf Prognosefreiheit<sup>31</sup> beruhende, stete Optionenvielfalt. Unter Optionen verstehen wir Denk- und Handlungsalternativen bezogen auf Untersuchungsgegenstände und die dazu passenden Bearbeitungen: Optionenvielfalt in der Linguistik bezieht sich auf ihre Untersuchungsgegenstände und Erkenntnisinteressen sowie auf ihre epistemologischen und methodologischen Praktiken. Dies bedeutet, die Linguistik für neue sprachliche Aspekte und für Denkstile anderer Denkkollektive zu öffnen, was unseres Erachtens keine Schwächung linguistischer Forschung ist. Es handelt sich um eine Transformation der Disziplin, die mit

24 Taleb, Nassim Nicholas: *Antifragilität. Anleitung für eine Welt, die wir nicht verstehen*, übers. von Susanne Held, München: btb Verlag 2014, S. 25.

25 Taleb: *The black swan*, S. xix; herv. i. Orig.

26 Vgl. Ebd., S. xviii.

27 Taleb, Nassim Nicholas: *Antifragile: Things That Gain from Disorder*, Random House Publishing Group 2012, S. 3.

28 Bei fragil, robust und antifragil handelt es sich um relative Begriffe vgl. Taleb: *Antifragilität*, S. 47.

29 Vgl. Ebd., S. 33–34.

30 Vgl. Ebd., S. 25.

31 Vgl. Ebd., S. 204.

der Abnahme von stark theoriegebundenem, disziplinbegründendem Forschen zugunsten eher flexibler ephemerer Forschung mit Theorien mittlerer Reichweite einhergeht.<sup>32</sup>

### 4.1 Untersuchungsgegenstände und Erkenntnisinteresse

Hier bedeutet Optionenvielfalt, die Veränderung des Untersuchungsgegenstands gedanklich zuzulassen und zu weiteren Sprachkonzepten des Denkens und Kommunizierens zu gelangen, um vielfältige neue Erkenntnisinteressen zu produzieren. So wie Umstände, die die freie Kommunikation aus politischen Gründen einschränken (z. B. Zensur, Diktatur) oder aus kognitiven Gründen Sprachbarrieren errichten (etwa in den Bereichen Aphasie, Leichte Sprache) zu neuen Forschungsfragen geführt haben, kann eine aufgeschlossene Linguistik im digitalen Wandel fortlaufend neue linguistisch relevante Dimensionen entdecken, z. B. Darstellungsformen von Sprache, Sprachsteuerung, Speicherung, Mensch-Maschinen-Kommunikation.

Wenn man davon ausgeht, dass das In-Beziehung-setzen älterer untersuchter Praktiken (z. B. Graffiti) und neuerer Praktiken (z. B. Trollen in Sozialen Medien) zu neuen Fragestellungen führt, ist dies ein Argument für eine Linguistik, die sich neuen sprachlichen Erscheinungsformen zuwendet, dabei aber die bisherigen Forschungsleistungen nicht vernachlässigt. Gerade die Kombination von Aspekten aus dem reichen Schatz des Faches (von Korpuslinguistik über Sprachkritik bis Rhetorik) und unerwartet neuen Aspekten kann zur Vervielfachung von Forschungsideen führen.

Ein weiteres Feld ergibt sich im Bereich der Forschung zur Metasprache und vor allem zur Metapragmatik: Viele Effekte medialen und auch digitalen Wandels und so gut wie jede neue sprachliche Handlung werden irgendwann von der Gesellschaft sprachlich aufgegriffen und häufig auch bewertet. In einem weiteren Sinn gilt dies für die Linguistik insgesamt, denn überspitzt formuliert könnte man sagen: Jede unerwartete Veränderung des Umstands, dass wir täglich Sprache produzieren, distribuieren, rezipieren und speichern, kann eine Bereicherung der Linguistik um einen neuen potenziellen Untersuchungsaspekt sein: Da wir in irgendeiner Weise auf das Unerwartete sprachlich reagieren werden.

### 4.2 Epistemologische und methodologische Praktiken

Ins Feld der epistemologischen und methodologischen Praktiken in der Linguistik fällt der in Abschnitt 2 skizzierte Konflikt zwischen linguistischen und alinguistischen Methoden zur maschinellen Analyse von Sprache. Die Taktik der kritischen Würdigung alinguistischer Methoden aus linguistischer Perspektive scheint uns zwar wichtig zu sein. Um aber die Linguistik antifragiler zu machen, sind weitere Strategien wichtig:

1) **Einverleibung** alinguistischer Methoden, wobei dies über die kritische Reflexion hinaus geht. Vielmehr geht es darum, solche Methoden zu verstehen, zu nutzen und dabei ins eigene Denkkollektiv zu integrieren. Es geht also nicht um die Abgrenzung gegenüber der anderen Disziplin und die Frage, welche Nische für die Linguistik bleibt, sondern darum, alinguistische Methoden selber zu praktizieren, dabei jedoch im Sinne eines *Hackings* und der *Bricolage*.<sup>33</sup>

<sup>32</sup> Vgl. Taleb: Antifragilität.

<sup>33</sup> Vgl. Ebd.



Der Begriff des Hackens stammt ursprünglich aus der Computerszene, wird dort aber als darüber hinaus gehende Ethik verstanden: «a philosophy of sharing, openness, decentralization, and getting your hands on machines at any cost to improve the machines and to improve the world».<sup>34</sup> Die pathetische Zielformulierung ist dabei weniger wichtig; worauf es ankommt, ist, nicht davor zurückzuschrecken, alinguistische Methoden auseinanderzunehmen, zu modifizieren – also zu basteln und zu experimentieren – und ihrem ursprünglichen Zweck zu entziehen. Das bedeutet auch, sich bewusst gegen methodische Kanons zu wenden (Bricolage), mit dem Ziel, bestehende Denkstile zu verändern oder neue ins Denkkollektiv zu integrieren.

**2) Vervielfältigung epistemologischer und methodologischer Optionen:** Die Einverleibung alinguistischer Methoden im Sinne des Hackings, der Bricolage und des Experiments führt im besten Fall zu einer Vervielfältigung des erkenntnistheoretischen und methodologischen Potenzials – und zu «serendipity», also dem «glücklichen Zufallsfund», ein Prinzip, auf das Schubert verweist.<sup>35</sup> Dies bedeutet: Mit jedem Ausflug disziplinärer Methoden, Fragestellungen oder Gegenstände in andere Disziplinen, wie etwa des Gegenstands Sprache in alinguistische Disziplinen, und wieder zurück, wird die Ursprungsdisziplin reicher. Dies deshalb, weil sich a) durch den Kontakt mit anderen Denkstilen die Methode, Fragestellung oder das Objekt anreichern mit neuen Aspekten. So etwa, weil neue Deutungsweisen, Zuspitzungen oder Modellierungen hinzukommen, im Falle der alinguistischen maschinellen Sprachverarbeitung etwa andere Modellierungen des Gegenstands Sprache. Jedoch auch, weil b) durch die Einverleibung in die eigene Disziplin erneut neue epistemologische oder methodologische Optionen entstehen.

Im Falle alinguistischer maschineller Textverarbeitung gewinnt die Linguistik a) neue Sprachvorstellungen (z. B. Sprache als Code, Relevanz von Metadaten) und b) veränderte Sprachpraktiken in der Gesellschaft, die untersucht werden können (z. B. Übersetzen, Recherchieren) und natürlich c) die alinguistische Methoden selber, ggf. eben auch in „verbastelter“ Form (z. B. Sentimentanalyse, algorithmusbasierte Rankings).

**3) Verschränkung digitaler und analoger Methoden:** Die Linguistik gewinnt zudem dann an Antifragilität, wenn die Verschränkung (und nicht das Nebeneinander) digitaler und analoger Praktiken gelingt. Also dann, wenn sowohl digital mit Papier und Stift als auch analog am Computer gearbeitet wird.

Ein Beispiel für diese Verschränkung ist eine diagrammatische Lesart computationeller Transformationen: Die heute üblichen Transformationen von Daten mittels Computer kann als klassische diagrammatische Operation<sup>36</sup> in diagrammatischen Grundfiguren zur Gewinnung von Erkenntnis aufgefasst werden, etwa der Erstellung von Listen durch Rekontextualisierungen von Fundorten.<sup>37</sup> Bei der (digitalen) korpuslinguistischen Konkordanzliste handelt es sich ja um die alte (analoge) Kulturtechnik der Erstellung eines Indexes, mit dem die Einheit der Originaltexte zerstört wird und dafür durch die Rekontextualisierung der Fundorte als Konkordanzliste ein neu zu deutender «Text» entstanden ist. Ein anderes Beispiel ist die diagramma-

34 Levy, Steven: *Hackers: Heroes of the Computer Revolution – 25th Anniversary Edition*, 1. Aufl., Sebastopol, CA: O'Reilly and Associates 2010, S. IX.

35 „Wie die Nomaden in Athen entdeckt wurden. Seltenes Nebeneinander zweier Wörter: Ein Beispiel für methodische Zufallsfundsuche in der Altertumskunde“, *FAZ*, 09.05.2018, S. N3.

36 Im Anschluss an Krämer, Sybille: *Figuration, Anschauung, Erkenntnis: Grundlinien einer Diagrammatologie*, Suhrkamp Verlag 2016.

37 Bubenhofer, Noah: „Visual Linguistics: Plädoyer für ein neues Forschungsfeld“, in: Bubenhofer, Noah und Marc Kupietz (Hrsg.): *Visualisierung sprachlicher Daten. Visual Linguistics, Praxis, und Tools*, Heidelberg: Heidelberg University Publishing 2018, S. 25–62.

tische Grundfigur der Partitur, mit der die Gleichzeitigkeit von Stimmen auf mehreren Ebenen dargestellt werden kann, und die als geisteswissenschaftlich-analoge Lesart von digitalen Auszeichnungssprachen aufgefasst werden kann.<sup>38</sup>

Solche Lesarten eröffnen neue Deutungshorizonte alinguistischer methodologischer Konzepte, so dass diese in der Linguistik zu neuen Forschungsfragen und Deutungshorizonten anstiften.

### 5. Fazit

Wichtig scheint uns, bei den Diskussionen zu den Veränderungen der Geisteswissenschaften durch den «black swan» Digitalität von einer Abwehrhaltung wegzukommen, die bloß Nischen für das eigene Fach sucht und sich abgrenzen will. Denn das sind Strategien, die die Robustheit des Faches im Blick haben: Das Fach soll möglichst so, wie es ist, bestehen bleiben können. Stattdessen trachtet die Denkfigur zur Antifragilität der Linguistik danach, Strategien zu formulieren, für die die Ungewissheit der Zukunft gerade von Vorteil ist, weil sie das Fach durch jegliche Veränderungen, seien sie auch noch so radikal, gestärkt hervorgehen lassen. Die genannten Eigenschaften und Strategien der Linguistik, die zu Teilen auf die Geisteswissenschaften generell übertragbar sind, erhöhen die Antifragilität der Disziplinen.

Interessanterweise führen diese Überlegungen in der Konsequenz zum gleichen Ergebnis wie etwa ein gegenüber den Digital Humanities kritisches, aus epistemologischen oder politischen Gründen auf Pluralität in den Wissenschaften pochendes Argument: Hermeneutik etwa als klassische geisteswissenschaftliche Methodologie oder Kritik als intellektuelle Haltung dürfe nicht geschwächt werden durch Förderung und Ausbau der Digital Humanities. Doch während das epistemologische Argument synchron gedacht ist, liegt dem strategischen Argument für eine antifragile Linguistik eine diachrone Überlegung zugrunde, die in der Diskussion um die Digital Humanities bisher zu kurz kommt. Gerade die Ungewissheit über die Zukunft sollte aber in einer Diskussion zur geisteswissenschaftlichen Disposition eine Rolle spielen.

---

38 Bubenhofer, Noah: „Visualisierungen in der Korpuslinguistik. Diagrammatische Operationen zur Gegenstandskonstitution, -Analyse und Ergebnispräsentation“, in: Kupietz, Marc und Thomas Schmidt (Hrsg.): Korpuslinguistik, Berlin / New York: De Gruyter 2018, S. 29.

## 6. Literaturverzeichnis

Andersen (2018): Chr. Andersen/ M. P. Ängsal/W. Czachur/ Ph. Dreesen/ U. Fix/ N. Kalwa,/ J. Kiesendahl/ J. Schiewe/ J. Spitzmüller/ B. Zimmermann, „Erkenntnis als soziale Praxis : Ludwik Flecks Wissenschaftstheorie aus sprachwissenschaftlicher Sicht“, *Denkstile in der deutschen Sprachwissenschaft : Bausteine einer Fachgeschichte aus dem Blickwinkel der Wissenschaftstheorie Ludwik Flecks* (2018), S. 11–65.

Bubenhofer (2018): N. Bubenhofer, „Visual Linguistics: Plädoyer für ein neues Forschungsfeld“, in: Bubenhofer, Noah und Marc Kupietz (Hrsg.): *Visualisierung sprachlicher Daten. Visual Linguistics, Praxis, und Tools*, Heidelberg: Heidelberg University Publishing 2018, S. 25–62.

Bubenhofer (2018): N. Bubenhofer, „Visualisierungen in der Korpuslinguistik. Diagrammatische Operationen zur Gegenstandskonstitution, -Analyse und Ergebnispräsentation“, in: Kupietz, Marc und Thomas Schmidt (Hrsg.): *Korpuslinguistik*, Berlin / New York: De Gruyter 2018.

Bubenhofer (2018): N. Bubenhofer, „Wenn ‚Linguistik‘ in ‚Korpuslinguistik‘ bedeutungslos wird. Vier Thesen zur Zukunft der Korpuslinguistik“, *OBST. Osnabrücker Beiträge zur Sprachtheorie* 1 (2018).

Dreesen (2018), Ph. Dreesen, „Denkstimmung der Skepsis : ein Beschreibungsansatz zur Entstehung der Diskurslinguistik“, *Denkstile in der deutschen Sprachwissenschaft : Bausteine einer Fachgeschichte aus dem Blickwinkel der Wissenschaftstheorie Ludwik Flecks* (2018), S. 167–190.

Fleck (2011): L. Fleck, „Das Problem einer Theorie des Erkennens“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011, S. 260–309.

Fleck (1980): L. Fleck, *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*, hg. von Lothar Schäfer und Thomas Schnelle, Frankfurt am Main: Suhrkamp 1980.

Fleck (2011): L. Fleck, „Schauen, sehen, wissen“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011.

Fleck (2011): L. Fleck, „Wie entstand die Bordet-Wassermann-Reaktion und wie entsteht eine wissenschaftliche Entdeckung im allgemeinen?“, in: Werner, Sylwia und Claus Zittel (Hrsg.): *Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse*, Frankfurt am Main: Suhrkamp 2011, S. 181–210.

Goodfellow (2016): I. Goodfellow/ Y. Bengio/A. Courville, *Deep Learning*, MIT Press 2016, <http://www.deeplearningbook.org>.

Kay (1977): A. Kay/ A. Goldberg: „Personal Dynamic Media“, *Computer* 10/3 (1977), S. 31–41.

Krämer (2016): S. Krämer, *Figuration, Anschauung, Erkenntnis: Grundlinien einer Diagrammatologie*, Suhrkamp Verlag 2016.

Lévy (1998): P. Lévy, „Die Erfindung des Computers“, in: Serres, Michel (Hrsg.): *Elemente einer Geschichte der Wissenschaften*, 2. Aufl., Frankfurt am Main: Suhrkamp Verlag 1998, S. 905–944.

Levy (2010): St. Levy, *Hackers: Heroes of the Computer Revolution – 25th Anniversary Edition*, 1. Aufl., Sebastopol, CA: O’Reilly and Associates 2010.

Olah (2018): Ch. Olah u. a., „The Building Blocks of Interpretability“, *Distill* (2018).

Schubert (2018): Ch. Schubert, „Wie die Nomaden in Athen entdeckt wurden. Seltenes Nebeneinander zweier Wörter: Ein Beispiel für methodische Zufallsfundsuche in der Altertumskunde“, *FAZ*, 09.05.2018, S. N3.

Stücheli-Herlach: P. Stücheli-Herlach/ M. Ehrensberger-Dow/ Ph. Dreesen, „Energiediskurse in der Schweiz: Anwendungsorientierte Forschung eines Kommunikationsfelds mittels digitaler Daten“, *Working Papers in Applied Linguistics* (im Druck).

Taleb (2012): N. N. Taleb, *Antifragile: Things That Gain from Disorder*, Random House Publishing Group 2012.

Taleb (2012): N. N. Taleb, *Antifragilität. Anleitung für eine Welt, die wir nicht verstehen*, übers. von Susanne Held, München: btb Verlag 2014.

Taleb (2012): N. N. Taleb, *The black swan: the impact of the highly improbable*, New York: Random House 2007.

Uszkoreit (2009): H. Uszkoreit, „Linguistics in Computational Linguistics: Observations and Predictions“Athens, Greece: Association for Computational Linguistics, 03.2009, <http://www.aclweb.org/anthology/W09-0105>.

## Autorenkontakt<sup>39</sup>

### **Dr. Noah Bubenhofer**

Zurich University of Applied Sciences ZHAW  
LCC, Digital Linguistics  
Theaterstr. 15c  
CH-8401 Winterthur

Email: [noah.bubenhofer@zhaw.ch](mailto:noah.bubenhofer@zhaw.ch)

### **Dr. Philipp Dreesen**

Zurich University of Applied Sciences ZHAW  
LCC, Digital Linguistics  
Theaterstr. 15c  
CH-8401 Winterthur

Email: [philipp.dreesen@zhaw.ch](mailto:philipp.dreesen@zhaw.ch)

---

<sup>39</sup> Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren. Alle Inhalte dieses Beitrages unterstehen, soweit nicht anders gekennzeichnet, der Lizenz CC BY 4.0.