

Paraphrase als Ähnlichkeitsbeziehung. Ein digitaler Zugang zu einem intertextuellen Phänomen

Kurt Sier / Eva Wöckener-Gade

Abstract Wer bereit ist, sich auf das (relativ) neue Forschungsparadigma der Digital Humanities einzulassen, sieht sich bisweilen genötigt, traditionelle geisteswissenschaftliche Kategorien und Begrifflichkeiten zu überdenken und gegebenenfalls anzupassen. Im vorliegenden Beitrag wird dies am Konzept der Paraphrase als Phänomen der Intertextualität illustriert. In einem ersten Teil soll es um den Hinweis gehen, dass Platons Reflexion auf das Problem der Ähnlichkeit im *tritos anthropos*-Argument des *Parmenides* dabei helfen kann, geisteswissenschaftliche und informatische Forschung ins Gespräch zu bringen. Im Anschluss wird an einem Beispiel aus Aristoteles' *Politika* vorgeführt, dass die digitale Messung von Ähnlichkeiten, hier mithilfe der Word Mover's Distance, unsere Sicht auf das Konzept der Intertextualität und auf das speziellere Phänomen der Paraphrase nicht unwesentlich bereichern kann.

Keywords Paraphrase, Ähnlichkeit, Intertextualität, Transformation, Word Mover's Distance, Word2Vec, *tritos anthropos*-Argument

Das Forschungsprojekt, über das die vorliegende Publikation Auskunft gibt, verfolgt einerseits ein sachliches, andererseits ein methodisches Anliegen. In der Sache geht es uns darum, einen Beitrag zur Aufarbeitung der Platon-Rezeption in der griechischen Antike zu leisten und durch den Einsatz von Verfahren der modernen Informationswissenschaften einem alten Forschungsdesiderat zu begegnen. Um Missverständnissen vorzubeugen, sei vorausgeschickt, dass es den an dem Projekt beteiligten Geisteswissenschaften völlig fern liegt, die traditionellen hermeneutischen Methoden in Frage stellen zu wollen. Natürlich gibt es keinen anderen Weg als die genaue Lektüre und Interpretation der Texte, wenn man die Ausstrahlung der platonischen Philosophie verstehend nachzuvollziehen sucht. Allerdings, was die Auffindung und Sammlung der Belege angeht, ist kaum zu bestreiten, dass die neuen Technologien dabei helfen können, zu vollständigeren und leichter zu gewinnenden Ergebnissen zu gelangen, als es der Philologie und Philosophie-

geschichte bisher möglich war. Unser Projekt will dies näher begründen und an Beispielen erläutern, aber es hat zugleich einen methodologisch-paradigmatischen Schwerpunkt, der auch für andere Bereiche der Altertumswissenschaften und der Digital Humanities von Interesse ist.

Das Thema des Projekts knüpft an eine frühere Kooperation von Informatikern und Geisteswissenschaftlern an der Universität Leipzig an, bei der die Gräzistik mit dem Teilprojekt vertreten war, die indirekte Platon-Überlieferung durch eine digitalisierte Suche nach Zitaten der platonischen Schriften in der späteren Literatur aufzuhellen. Eine entsprechende Recherche-Maske steht im Internet zur Verfügung (www.eaqua.net). Bereits vor zehn Jahren war indes klar, dass wörtliche Zitate ja nur einen begrenzten Teil der antiken Platon-Rezeption ausmachen und die eigentliche Herausforderung darin bestehen würde, auch solche Bezugnahmen zu identifizieren, die auf den platonischen Text nur anspielen oder ihn paraphrasieren. Es ist leicht zu sehen, dass dies methodologisch auf eine wesentlich anspruchsvollere Fragestellung als im Falle der Zitate hinausläuft. Denn hier fehlt mit der sprachlichen Übereinstimmung das Substrat, auf das eine automatisierte Recherche sich beziehen ließe. Wo der rezipierende Text vom rezipierten im Wortlaut abweicht, liegt die eine Seite der Relation zunächst im Dunkeln, und leider verhält es sich mit den antiken Texten anders als mit heutigen Produkten, bei denen die informatische Linguistik sich erfolgreich bemüht, ein transponierendes Verstehen von Vokabular und Grammatik auch künstlichen Systemen anzutrainieren. Davon sind wir im Umgang mit der antiken Literatur noch weit entfernt. Gleichwohl hat sich bei unserem Projekt eine fruchtbare Zusammenarbeit der geisteswissenschaftlichen und der informatischen Disziplinen ergeben, deren Resultate, wie wir meinen, ein Modell für eine operationalisierbare Paraphrasen-Suche auch in antiken Texten an die Hand geben.

Im vorliegenden Beitrag soll es um zweierlei gehen. In der ersten Hälfte schlagen wir eine Konzeptualisierung des Intertextualitäts- und Paraphrasenbegriffs vor, die es möglich macht, nicht-wörtliche Textbeziehungen in einer Weise zu beschreiben, die mit den Methoden der Informatik kompatibel ist. Die zweite Hälfte beleuchtet unseren Ansatz auf dem Hintergrund der gegenwärtigen Intertextualitätsforschung und illustriert an einem Beispiel, dass und wie die Einbeziehung informatischer Methoden zu neuen Erkenntnissen verhelfen kann.

I

Wie Heidegger sagte,¹ sind wir „nicht, und wenn, dann nur selten und dabei kaum, in der Lage, eine Beziehung, die zwischen zwei Dingen, zwischen zwei Wesen wal tet, rein aus ihr selbst zu erfahren. Wir stellen uns die Beziehung sogleich von dem aus vor, was jeweils in der Beziehung steht. Wir sind wenig darüber verständigt, wodurch und woher sich die Beziehung ergibt und wie sie als diese Beziehung ist“. Das Projekt einer operationalisierbaren Suche nach Platon-Paraphrasen gibt Anlass, auch Heideggers Frage zu überdenken. Denn was wir intendieren, ist ja, Textrela tionen zu finden, bei denen der eine Part nicht vorgegeben ist, sondern gerade erst eruiert werden soll, so dass die Form der herzustellen Relation selbst genauer in den Blick zu nehmen ist. Wir lassen uns dabei von zwei Vorüberlegungen leiten.²

(1) Die bei der Paraphrase gegebene Relation ist eine Spezies der *Intertextuali tät*. Unter ‚Paraphrase‘ verstehen wir

(a) den Fall, dass ein Text B auf einen Text A in der Weise bezogen ist, dass er mit anderen Mitteln entweder den gleichen oder einen konträr entgegengesetzten (oder in signifikanter Weise abweichenden) Sachverhalt zum Ausdruck bringt oder darstellt;

(b) den Fall, dass ein Text B nicht unmittelbar auf einen Text A bezogen ist, sondern auf einen übergreifenden Kontext verweist. Auf die Platon-Rezeption angewandt: Text B bezieht sich nicht auf eine bestimmte Stelle in den platonischen Dialogen, sondern auf den jeweiligen Dialog oder spezifische Konzepte/Formulierungen des *Corpus Platonicum* im ganzen.

Die Unterscheidung zwischen (a) und (b) – wir sprechen von Paraphrasen im ‚engeren‘ und im ‚weiteren‘ Sinn – ist nicht immer sicher, so dass auch (b) in Über legungen zum Paraphrasenbegriff einzubeziehen wäre. Unser primäres Interesse gilt jedoch der Relation (a), und wir rechnen hier mit unterschiedlichen Graden der Übereinstimmung. Text B nimmt zwar nicht eins zu eins und wörtlich auf Text A Bezug (sonst wäre er ein Zitat), aber die vorhandenen Abweichungen müssen sich als Transformationen von A verstehen lassen, wenn B als Paraphrase gelten soll, und die Kriterien für die Identifizierung einer Transformation variieren von Fall zu Fall und sind teils mehr, teils weniger valid. Das führt auf das für das Projekt zwar sehr relevante, aber in seinem zeitlichen Rahmen nicht lösbare Problem, wie diese Kriterien selbst adäquat zu bestimmen wären. Das Vorhandensein einer themati schen Konvergenz und die semantische oder sprachlich-syntaktische Austauschbarkeit von Textbausteinen stellen gewiss eine notwendige, aber noch keine hinreichende Bedingung für die Annahme einer Paraphrase im Einzelfall dar. Unser

1 (1959) 188.

2 Vgl. auch den Forschungsüberblick im [zweiten Teil dieses Beitrags](#).

Projekt muss diese methodologische Frage einstweilen offen lassen; es versteht sich als Handreichung für die Interpretation der in ihm aufgewiesenen ‚Parallelen‘ und beansprucht nicht, Gründe für die Annahme zu liefern, dass es sich bei den Parallelen um literarische Beziehungen handelt.³

(2) Intertextualität ist ein *Rezeptionsphänomen*. Auch wo Text B evidente Signale für seine Bezogenheit auf Text A aufweist (wie z. B. bei Zitaten oder bei einem Verhältnis wie dem von Vergils *Aeneis* zu Homer), handelt es sich zunächst um eine Relation in *eine* Richtung (im Sinne von *agens* und *patiens*), und es bleibt dem Rezipienten vorbehalten, die Texte miteinander ins Gespräch zu bringen und ihre Beziehung als ein reziprokes Verhältnis zu verstehen. Beim vorliegenden Projekt ist der jeweilige Text B in der Regel eine Unbekannte, etwas Aufzufindendes, womit die hermeneutische Herausforderung steigt. Denn es gilt, zuerst die Bedingungen zu beschreiben, unter denen sich ein Text x als ein Text B verstehen und mit Text A in einen Dialog bringen lässt. Die an dem Projekt beteiligten Disziplinen stehen hier für tendenziell unterschiedliche Ansätze. In den Geisteswissenschaften werden intertextuelle Beziehungen meist als Verhältnis zwischen Prätext und Posttext rekonstruiert, d. h. im Sinn einer vertikalen Struktur, bei der Text B von Text A ‚abhängt‘. Für unser Projekt, das die ‚B‘-Texte erst noch erschließen will, ist dieses Konzept untauglich. Wir tendieren, im Anschluss an Julia Kristeva (1972) und andere, eher zur Annahme einer offenen, horizontalen Struktur, in der intertextuelle Relationen sich als ein Netzwerk gleichberechtigter Kommunikationspartner darstellen. Allerdings kann dies für uns nur eine vorläufige Orientierung sein. Denn für das Vorhaben, Platon-Paraphrasen zu identifizieren, hilft es wenig, sozusagen die gesamte griechische Literatur als ‚Kontext‘ zu verstehen, da wir ja auf spezifische, unter dem Aspekt der Ähnlichkeit signifikante Kontexte hinauswollen. Die Suche nach den B-Texten kommt ohne die Annahme der Dependenz und einer transitiv-hierarchischen Beziehung nicht aus, und die Frage ist vielmehr, wie der Ausgangspunkt sich methodisch schlüssig bestimmen lässt, wenn sein Korrelat unbekannt ist.

Vielleicht könnte man die *ratio* unserer Methode so beschreiben: wir werfen den jeweiligen Platon-Text wie einen Anker aus und versuchen herauszufinden, was an ähnlichen Texten sich in seinem Umkreis sammelt. Für die Formalisierung und Automatisierung dieses Vorgehens sind die informatischen Disziplinen zuständig, und sie haben im Kontext des Projekts in der Tat überraschende und wegweisende Ergebnisse erzielt, indem sie Textbeziehungen als gerichtete Bewegungen von A nach x/B beschreib- und analysierbar gemacht haben – u. a. durch das Word2Vec-Modell, das auf der Annahme beruht, dass Wörter mit ähnlicher Bedeutung

3 Allerdings rechnen wir, wie gesagt, mit Gradunterschieden bei den ermittelten Textähnlichkeiten, die sich wohl auch informatisch beschreiben lassen – z. B. in einem Algorithmus der Form: je weniger eine Text-Relation durch den Kontext angezeigt wird, umso mehr sprachlicher Aufwand ist nötig, um die Bezugnahme deutlich zu machen.

sich in ähnlichen Kontexten finden,⁴ und das Ähnlichkeit als Distanz zwischen Wortvektoren messbar werden lässt. Den geisteswissenschaftlichen Fächern fällt bei der Kooperation die Aufgabe zu, genauer zu erklären, wie der von der Informatik vorausgesetzte Begriff der Ähnlichkeit zu definieren ist, wenn er heuristisch fruchtbar sein soll. Das ist gewiss noch ausbaufähig, aber unser Ansatz stellt sich derzeit wie folgt dar:

Auch die bisherige Intertextualitätsforschung geht in der Regel von der Annahme aus, dass die Relation zwischen zwei Texten ein Drittes voraussetzt, das den beiden Texten gemeinsam zukommt. Indes sind die Bestimmungen dieses Dritten vielfach vage und auch zirkulär, insofern sie sich auf etwas beziehen, das erst noch zu erweisen wäre – z. B. eine Übereinstimmung der Texte in Bezug auf ihr Thema, ihre Sinnrichtung oder ihre Konzepte. Die Semantik ist ein unsicherer Wegweiser, wenn man von Text A zu Text B gelangen will, aber sie ist als Kriterium natürlich unverzichtbar. Wie wäre ihre einheitsstiftende Funktion in eine methodisch valide Form zu bringen? Vielleicht, so unser Modell, indem man das vermittelnde ‚Dritte‘ nicht direkt in der Semantik, sondern in einem gemeinsamen Bezug der jeweiligen Texte auf einen ‚Hypertext‘ findet. Was das bedeutet, sei kurz skizziert.

Uns scheint, dass die Besinnung auf Platon hier in der Sache und einer durchaus nicht anachronistischen Weise weiterhelfen kann. Karl Popper hat schön ausgeführt, wie Platons Ideenkonzeption die erste und eine in der Struktur nach wie vor gültige Erklärung des Phänomens der Ähnlichkeit liefert.⁵ Und in der Tat, wenn man nach einem übergreifenden Feature sucht, das die gesamte platonische Philosophie von Anfang bis Ende bestimmt, wäre es die Frage nach den Formen der Relation und

4 Diese sogenannte ‚distributionelle Hypothese‘ wurde von Harris (1957) aufgestellt.

5 (1993) 203: „Wir müssen die eng mit dem Animismus verknüpfte (und für Aristoteles im Gegensatz zu Platon charakteristische) Anschauung aufgeben, daß man sich auf die jedem individuellen oder einzelnen Ding inhärenten wesentlichen Eigenschaften berufen kann, um das Verhalten dieses Dinges zu erklären. Denn dieser Anschauung gelingt es durchaus nicht, Licht auf das Problem zu werfen, warum verschiedene individuelle Dinge sich auf ähnliche Weise verhalten. Wenn gesagt wird: ‚Weil ihre wesentlichen Eigenschaften ähnlich sind‘, so erhebt sich die neue Frage, warum es nicht ebenso viele wesentliche Eigenschaften geben sollte, wie es verschiedene Dinge gibt. Platon versuchte genau dieses Problem zu lösen, indem er sagte, daß ähnliche individuelle Dinge Abkömmlinge, und daher Abbilder, derselben ursprünglichen ‚Form‘ sind, die ‚außerhalb‘ der verschiedenen individuellen Dinge besteht und ‚älter‘ als diese und ihnen ‚überlegen‘ ist; und in der Tat, wir haben bis jetzt keine bessere Theorie der Ähnlichkeit. Selbst heutzutage berufen wir uns auf ihren gemeinsamen Ursprung, wenn wir die Ähnlichkeit zweier Menschen oder [...] erklären wollen; das heißt, wir erklären Ähnlichkeit hauptsächlich genetisch; und wenn wir daraus ein metaphysisches System machen, so neigt es dazu, eine historizistische Philosophie zu werden. Platons Lösung wurde von Aristoteles verworfen; aber da Aristoteles‘ Fassung des Essentialismus nicht einmal den Hinweis einer Lösung enthält, so scheint es, daß er niemals das Problem erfaßt hat.“

Vermittlung, die die Differenz der individuellen Dinge überbrückt und uns einen synoptisch-ordnenden Blick auf die Wirklichkeit ermöglicht. Charakteristisch z. B. das Programm im *Timaios* (31 b): ‚Dass zwei Dinge allein, ohne ein Drittes, auf schöne Weise zusammengehen, ist unmöglich; denn es muss in der Mitte zwischen beiden ein Band geben, das sie zusammenführt. Das schönste Band aber ist dasjenige, das sich selbst und das zu Verbindende am meisten zu einer Einheit macht, und dies am schönsten zu leisten ist die Proportion (*analogia*) geeignet‘. Auch die Ideenkonzeption in ihren verschiedenen Spielarten hat u. a., wie Popper sagt, die Funktion, das vergleichende Verbinden und Trennen der Phänomene zu begründen und eine Ähnlichkeit des Differenten zu erklären. Platon rechnet mit einer dreistelligen Relation: die Ähnlichkeit zwischen den vergleichbaren Dingen beruht auf der *methexis*, ihrer ‚Teilhabe‘ an der Form oder Idee, und der gemeinsame Bezug auf das *paradeigma* erklärt ursächlich die Ähnlichkeitsbeziehung zwischen den Dingen. Man muss kein Anhänger seiner Metaphysik sein, um den heuristischen Wert dieses Ansatzes zu erkennen, der das semantische Kriterium für das vergleichende Ordnen der Dinge nicht in deren individuell-kontingenter Beschaffenheit, sondern in einem vorausliegenden Sachgehalt ansiedelt, der dem Zweifel entzogen ist. Es scheint nicht abwegig, Platons Konzeption (*mutatis mutandis*) als Modell auch für die Auffindung oder Begründung intertextueller Beziehungen heranzuziehen und ihn beim Projekt der Paraphrasen-Suche zum Interpretieren seiner eigenen Rezeption zu machen.

Nun kann dies allerdings nicht so aussehen, dass der Text der platonischen Dialoge quasi als ideales Paradeigma und die gesuchten Paraphrasen als seine Instantiierungen zu betrachten wären. Denn zum einen ist der platonische Text selbst ein individuelles Element der von ihm begründeten Tradition des Platonismus und zum anderen ergäbe seine Absolut-Setzung höchstens den theoretischen Befund, dass die auf ihn bezogenen Paraphrasen einander ähnlich sein müssen, aber noch kein Kriterium dafür, dass und inwiefern die betreffenden Texte tatsächlich als *Paraphrasen* von ihm zu interpretieren sind. Weiterhelfen kann hier das (von Aristoteles so genannte) Argument des ‚dritten Menschen‘ (*tritos anthropos*), das im Dialog *Parmenides* diskutiert wird. Dabei handelt es sich um einen *Einwand* gegen die Ideenlehre, und die Frage der Validität dessen, was Parmenides gegen Sokrates vorbringt, ist in der Forschung sehr umstritten.

Kurz gesagt, geht es darum, dass die Ideenannahme in einen unendlichen Regress führe, da die Ähnlichkeitsbeziehung zwischen der Idee und den an ihr teilhabenden Dingen eine weitere Idee als Drittes fordere, das die Annahme der Ähnlichkeit begründe. Die Argumentation vollzieht sich in zwei Schritten. (1: 132 a b) ‚Wenn dir eine Vielzahl von Dingen groß erscheint, so gewinnst du im Blick auf die Gesamtheit vielleicht den Eindruck, es gebe *eine* identische Form, weshalb du meinst, das Große sei *eines*. – Du hast recht. – Wie aber verhält es sich mit dem Großen selbst und den anderen großen Dingen? Wenn du in gleicher Weise mit der Seele

auf die Gesamtheit blickst, wird dann nicht aufs neue irgend *ein* Großes auftauchen, dem es sich verdankt, dass all dies groß erscheint? – So sieht es aus. Wie es scheint, ergibt sich das von Parmenides behauptete Problem daraus, dass er die Relation der Dinge zur Idee überspringt, d. h. die Eigenschaft, durch die die Dinge auf die Idee verweisen, mit dieser gleichsetzt. Das ist nicht das, was die Ideenkonzeption eigentlich meint, und es droht auch kein endloser Regress, da die Eigenschaft eben das ist, wodurch die Dinge an der Idee teilhaben, und die Ähnlichkeit des Teilhabenden mit dem, woran es teilhat, durch die asymmetrische Relation der *methexis* restlos erklärt wird – die Annahme einer Über-Idee erübrigt sich. (z: 132 c–133 a) Während Sokrates die Teilhabe als Abbild- und Ähnlichkeitsbeziehung der Abbildungen zum Urbild bezeichnet, macht Parmenides daraus eine symmetrische Relation und folgert, da das Paradeigma seinerseits den ihm ähnlichen Dingen ähnlich sei, müsse es an derselben Idee wie diese teilhaben; und da diese dann ebenfalls in einem Ähnlichkeitsverhältnis stehe, ergebe sich ein unendlicher Regress. Doch die Teilhabe meint *per definitionem* eine nichtsymmetrische Relation, und Abbild und Urbild mögen allenfalls, in unterschiedlicher Weise, an der Idee des Ähnlichen teilhaben, nicht aber an einer ähnlichen Idee. „There may be many similar photographs of the same person [...]; but you cannot argue that the person cannot be like the photographs unless he is himself a picture of a second person.“⁶ Im ganzen erscheint das *tritos anthropos*-Argument im Sinne Platons nicht als valider Einwand gegen die Ideenkonzeption, sondern ist wohl eher als ein Ausschnitt der in der platonischen Akademie geführten Diskussionen zu betrachten. Gleichwohl scheint die Argumentation hilfreich, wenn es um die Struktur des Verhältnisses von empirischen Texten geht.

Denn man kann von ihr aus ein Intertextualitätsmodell entwerfen, das den jeweiligen Bezugstext als Abspiegelung eines ‚idealen‘ Texts versteht, der die Grundlage für die Annahme bildet, dass der vermeintliche Bezugstext und der Text oder die Texte, den/die wir auf ihn beziehen, einen gemeinsamen ‚Ursprung‘ und Konvergenzpunkt haben. Auf die Frage der Platon-Paraphrasen angewandt: als Paradeigma ist (sozusagen) die ‚Idee‘ des jeweiligen Platon-Texts anzusetzen, und die Beziehung zwischen den Texten ist anhand des Kriteriums der Nähe zu diesem ‚Vorbild‘ zu beurteilen. Der empirische Platon-Text wäre mit dem Paradeigma identisch, während die übrigen Texte an dem Abstand von diesem zu messen wären. Das ist sicherlich nur ein erster Schritt in Richtung auf eine Objektivierung des Intertextualitätskonzepts und eine ‚Theorie‘ der Paraphrase, die geisteswissenschaftliche und informatische Ansätze zusammenführt. Jedenfalls hat die Kooperation der Disziplinen bei den Geisteswissenschaftlern zu einer neuen Wahrnehmung der Defizite ihrer Erklärungsmodelle und zur Suche nach einem begründungsfähigeren Neuansatz geführt.

6 Cornford (1957) 94.

II

Der oben entwickelte theoretische Ansatz zum Verständnis von Paraphrasen als intertextuellen Phänomenen soll im Folgenden konkretisiert und auf seine Anwendbarkeit geprüft werden. Dafür wird er zunächst in das Spannungsfeld bisheriger Forschung zum Thema eingeordnet, dann werden an einem Beispiel die Operationalisierbarkeit mithilfe von Distanzmaßen und die daraus resultierenden Vorteile für die Textanalyse vorgestellt.

Um die genaue Bestimmung der beiden hier relevanten Konzepte, der Intertextualität und der Paraphrase, wird in der Forschung seit längerem gerungen, wobei ein Konsens nur für ein sehr weit gefasstes Verständnis der Begriffe gilt⁷ und vor allem die Operationalisierbarkeit als zentrales Problem der verschiedenen Ansätze ausgemacht wurde.⁸ Verantwortlich für die Probleme ist u. a. der Umstand, dass die genauere Bestimmung der Phänomene auf Kriterien fußt, die von der Beurteilung durch den jeweiligen Anwender abhängen. Evident ist dies für die Intertextualität als Rezeptionsphänomen: Die intertextuelle Beziehung zwischen Post- und Prätext wird jeweils erst vom Rezipienten durch seine individuelle Interpretationsleistung hergestellt;⁹ das gleiche gilt im Prinzip auch für das Phänomen

7 Für die Intertextualität trifft dies in besonderem Maße zu; einschlägig ist immer noch die allgemeine Bestimmung von Pfister (1985a) 11: „Die Theorie der Intertextualität ist die Theorie der Beziehungen zwischen Texten. Dies ist unumstritten; umstritten jedoch ist, welche Arten von Beziehungen darunter subsumiert werden sollen. Und je nachdem, wieviel man darunter subsumiert, erscheint Intertextualität entweder als eine Eigenschaft von Texten allgemein oder als eine spezifische Eigenschaft bestimmter Texte oder Textklassen.“ Als Definition der Paraphrase wird oft die weiteren Spielraum lassende „approximate conceptual equivalence among outwardly different material“ von DeBeaugrande/Dressler (1981) 50 herangezogen, während viele engere Definitionen von semantischer Übereinstimmung ausgehen. Vgl. den Überblick bei Ho et al. (2012) 855 ff., die ebenfalls für eine weitere Definition plädieren („different words, phrases or sentences that are having similar meanings“ (853) bzw. „that express the same or almost the same meaning“ (858).

8 Vgl. für die Intertextualität Holthuis (1993) 22 ff. sowie Pfister (1985a) 18 und 25 sowie für die Paraphrase u. a. Ho et al. (2012) 852 ff. Die Schwierigkeit der Operationalisierbarkeit stellte sich in unserem Projekt in umso höherem Maße, als unsere Bestimmung von Intertextualität und Paraphrase nicht nur beansprucht, zur Analyse bekannter Instanzen anwendbar zu sein, sondern auch zur Suche nach bisher noch nicht entdeckten. Zudem sind die Texte unseres Forschungsobjekts – die Texte des *Thesaurus Linguae Graecae* (TLG-E) – sprachlich wie literarisch sehr heterogen, und diese Heterogenität bildet sich natürlich auch in den sprachlichen Einzelphänomenen ab. Beiden Umständen Rechnung tragend beruht daher unsere Theoriebildung nicht rein auf analytisch-deskriptiven, sondern auch auf explorativen Verfahren.

9 Vgl. Holthuis (1993) 31 ff., die sich im Anschluss an frühere Ansätze überzeugend gegen die Position wendet, Intertextualität als textimmanentes und vom Rezipienten entkoppelbares Phänomen zu fassen.

der Paraphrase, die ohne Zuordnung zum Prätext eben lediglich eine Phrase ist. Die meisten anwendungsorientierten Ansätze gehen davon aus, dass sich der Rezipient bei der Herstellung eines intertextuellen Bezugs von Signalen, die im Text feststellbar sind, leiten lässt;¹⁰ sie unterscheiden hier verschiedene Typen und machen diese zu Kriterien für die Analyse von intertextuellen Phänomenen. Einen wichtigen Beitrag auf diesem Gebiet hat Manfred Pfister geleistet. In seiner grundlegenden Arbeit zur Intertextualität hat er den Versuch unternommen, „Grade der Intensität von Intertextualität“ anhand verschiedener Kriterien differenzierbar zu machen, und vorgeschlagen, diesen (metaphorisch zu verstehende) mathematische Werte zuzuordnen, um eine Messbarkeit und damit Vergleichbarkeit verschiedener Instanzen zu erreichen.¹¹ So wertvoll Pfisters Kriterienkatalog und ähnliche jüngere Ansätze als Methode der Textanalyse erscheinen, bleiben diese doch in Vielem Annahmen verhaftet, die dem Ziel der Messbarkeit zuwider laufen: Um möglichst viele Signale im Text festmachen zu können, folgt Pfister einem eher strukturalistischen Ansatz und geht u. a. von bewusster Intention des Autors als Indikator von starker Intertextualität aus.¹² Letzteres ist gleich in mehrerer Hinsicht problematisch: Müsste eine solche Intention, wenn sie vorliegt, sich auch im Text manifestieren? Wenn dies der Fall wäre, könnte das entsprechende Signal dann auch vom Rezipienten/ Forscher zuverlässig aufgefunden und gedeutet werden? Und schließlich: Wie kann man sich eine Messung der Stärke dieses Signals vorstellen und wären aus einer solchen wiederum Rückschlüsse auf die Stärke der Intention zu ziehen?¹³ Aufgrund dieser und anderer Unsicherheiten erscheint die Annahme von Intention als ein Kriterium, dessen Bestimmung mit zu vielen Unbekannten arbeitet.¹⁴ Wahrscheinlich ist Intertextualität als globales Phänomen, wie

10 Nimmt man solche nicht an, müsste man schlichtweg eingestehen, dass Intertextualität nicht analysierbar sei, vgl. Holthuis (1993) 25 mit Bezug auf Lachmann (1990).

11 (1985a) 25–30.

12 (1985a) 27 in Abgrenzung zu „nur durch den Rezipienten willkürlich an den Text herangetragene[n] Prätexte(n)“. Noch stärker wird die Bedeutung von Intention für das Wesen der Intertextualität von Hermann (2007) 24 hervorgehoben, die eine rezeptionsorientierte Konzeption ablehnt.

13 Als „harte[n] Kern maximaler Intensität“ nennt Pfister ebd. den Fall eines bewussten und vom Autor eindeutig markierten Bezugs, der beim Rezipienten Kenntnis des Prätexts voraussetze. Schon eine solche Annahme erfordert viel Ausdeutung (z. B., ob der Autor beim Publikum denselben Wissensstand erwartet oder es durch die Markierung erst zur Lektüre des angesprochenen Werkes ermuntern möchte). Das gilt umso mehr für die häufigeren, interessanteren, weniger eindeutigen Fälle.

14 Hinzu kommt u. a. für unser Textkorpus, dass dies aufgrund großer Lücken in der Überlieferung nur einen kleinen Teil der literarischen Tradition der Antike mit teils unsicheren Datierungen abbildet. Diese Leerstellen verbieten in vielen Fällen, Abhängigkeiten zu postulieren.

Pfister selbst einschränkt, schlicht nicht messbar,¹⁵ da sie stets an den Rezipienten gekoppelt bleibt, auf den selbst strukturalistische Ansätze spätestens dann zurückfallen, wenn sie von individuellen Forschern unterschiedlich umgesetzt werden. Die Ergebnisse solcher Analysen sind daher zwangsläufig weder konsistent noch vergleichbar.

Günstiger scheint es zunächst bei dem klarer einzugrenzenden Phänomen der Paraphrase auszusehen. Ausgehend von den genannten sehr allgemeinen Definitionen wird der Prozess des Paraphrasierens oft als Transformation eines Prätextes in einen Posttext verstanden. Vor allem im Bereich der Linguistik wurden verschiedene Versuche unternommen, durch Analyse großer Mengen von Paraphrasen verschiedene Klassen und Typen lexikalischer, semantischer und syntaktischer Transformationen auszumachen und so zu bestimmen, was (und was nicht) Paraphrase ist.¹⁶ Die Stärken dieses eher deskriptiven als explikativen Ansatzes liegen, ähnlich wie bei Pfister, vor allem darin, dass man mit ihm Paraphrasen genauer beschreiben, kategorisieren und analysieren kann. Bei dem Versuch, solche Systematisierungen anzuwenden, stößt man jedoch schnell wieder auf Probleme. So gehen z. B. alle Ansätze von Auslassung/Hinzufügung als einer Form von Transformation aus;¹⁷ dabei müsste allerdings einerseits näher bestimmt werden, was unter diese fällt,¹⁸ und andererseits ist das, was ausgelassen bzw. hinzugefügt ist, nicht der eigentlichen Paraphrase zuzurechnen und taugt daher nur bedingt zu ihrer Charakterisierung. Zudem ist teils die Abgrenzung von anderen Phänomenen wie der Ersetzung schwierig. So listen Bhagat/Hovy als Beispiel für „Metaphor substitution“ (und nicht für Addition) *fog* ↔ *wall of fog*,¹⁹ Vila et al. klassifizieren dasselbe Beispiel als „Synthetic/analytic substitutions: Specifier addition/deletion“,²⁰ das ähnlich gelagerte Beispiel „a) A teaspoonful of vanilla b) Very little“, hingegen als „Same-polarity substitution“.²¹ Verschiedentlich wurde darauf hingewiesen, dass

15 (1985a) 30.

16 Vgl. Vila et al. (2014) 208 f. für einen Überblick über verschiedene Typologien sowie 211 f. für die von ihnen entwickelte, die frühere Ansätze zusammenfasst und das Phänomen abzugrenzen sucht. Noch nicht berücksichtigt ist dort die von Bhagat/Hovy (2013) vorgeschlagene Typologie. Trotz abweichender Ansätze listen die meisten Arbeiten 24–25 Typen auf, teils mit zusätzlich ausdifferenzierten Unterkategorien. Dronsch (2007) 37 f. hat vorgeschlagen, auch Intertextualität als „Transformation“ zu verstehen, verweist aber selbst auf die Schwierigkeit die einzelnen „Transformationsprozesse näher bestimmen zu können“.

17 Z. B. benennen Vila et al. (2014) 211 dies als eins der häufigsten Phänomene.

18 Das Paraphrasieren beginnt wie das Zitieren in der Regel mit der Auswahl eines oder mehrerer Textsegmente, die dann in einen neuen Kontext überführt werden, ist also per se selektiv. Diese Dekontextualisierung wäre vom Phänomen der Auslassung klar abzugrenzen.

19 (2013) 467.

20 (2014) 213.

21 (2014) 211.

die Kategorisierung von Paraphrasen je nach Bearbeiter der Datenerhebung stark variiert und versucht, dem mit möglichst engen Annotationsrichtlinien zu begegnen.²² Versteht man Paraphrase allerdings als Erscheinungsform der Intertextualität und damit als Rezeptionsphänomen im oben ausgeführten Sinn, wird deutlich, dass sie sich einer solchen eindeutigen Eingrenzung entzieht.²³ Vielmehr ist Vagheit als ein konstituierendes Merkmal von Paraphrasen zu verstehen.²⁴ Daher schlagen wir wie gesagt vor, die vielen Definitionsansätzen zugrundeliegende ‚Ähnlichkeit‘ zum Kriterium der Kategorisierung von Paraphrasen zu machen, die sich anhand einer (z.B. mathematischen) Abstraktionsebene bestimmen lässt und sich in verschiedenen Erscheinungsformen manifestieren kann. Wie genau sich unser Ansatz darstellt und welche Vorteile er gegenüber anderen bietet, soll im Folgenden an einem Beispiel verdeutlicht werden. Die prinzipielle Problemstellung, was eine Paraphrase eigentlich ausmacht und wie sie charakterisiert werden kann, lässt sich mit einer vergleichenden Analyse zweier auf Platon bezogener Aristoteles-Paraphrasen gut veranschaulichen.

III

Im zweiten Buch seiner *Politika* setzt sich Aristoteles mit den Staatstheorien seiner Vordenker auseinander, beginnend mit Platon und dessen *Politeia*. Erster Diskussionspunkt ist hier die Gütergemeinschaft, genauer das Problem, wie weit sie sich erstrecken soll. Nachdem Aristoteles die Fragestellung kurz umrissen hat, gibt er Platons Position folgendermaßen wieder (1261 a1–10):²⁵

ἀλλὰ πότερον ὅσων ἐνδέχεται κοινωνῆσαι, πάντων βέλτιον κοινωνεῖν τὴν μέλλουσαν οἰκίσεσθαι πόλιν καλῶς, ἢ τινῶν μὲν τινῶν δ' οὐ βέλτιον; ἐνδέχεται γὰρ καὶ τέκνων καὶ γυναικῶν καὶ κτημάτων κοινωνεῖν τοὺς πολίτας ἀλλήλοις, ὥσπερ ἐν τῇ Πολιτείᾳ τῇ Πλάτωνος· ἐκεῖ γὰρ ὁ Σωκράτης φησὶ δεῖν κοινὰ τὰ τέκνα καὶ τὰς γυναῖκας εἶναι καὶ τὰς κτήσεις. τοῦτο δὲ πότερον ὧς νῦν οὕτω βέλτιον ἔχει, ἢ κατὰ τὸν ἐν τῇ Πολιτείᾳ γεγραμμένον νόμον;

22 Vgl. u.a. Ho et al. (2012) 558 mit Verweis auf Herrera et al. (2007).

23 Vergleiche die zwar auf die Rede bezogene, aber auch für literarische Formen gültige Kritik von Ungeheuer (1972) 184f., Synonymie an der lexikalischen Ebene anzubinden und Paraphrasen rein darüber zu definieren, vernachlässige die Rolle des Individuums beim Prozess des Paraphrasierens.

24 Vgl. grundlegend zum Phänomen Pinkal (1985).

25 Hervorhebungen (kursiv) stammen i.F. von der Verfasserin.

Aber ist es besser, dass ein Staat, der gut eingerichtet sein soll, alles, was man gemeinschaftlich haben kann, gemeinschaftlich hat, oder ist es besser, dass er einiges gemeinschaftlich hat, einiges aber nicht? Es ist nämlich schon möglich, *dass die Bürger untereinander Nachkommen, Frauen und Besitz gemeinschaftlich haben*, wie in Platons *Politeia*; denn dort behauptet Sokrates, *dass die Nachkommen, die Frauen und die Besitztümer gemeinschaftlich sein müssen*. Ist es nun besser so, wie es sich jetzt verhält, oder so, wie es das in der *Politeia* vorgeschlagene Gesetz bestimmt? (Übers. nach Schwarz)

Aristoteles hat seine Bezugnahme für antike Gepflogenheiten sehr sorgfältig gekennzeichnet: Er verdeutlicht, dass er fremde Positionen diskutiert (dies bereits in 1260 b27 ff.)²⁶ und erwähnt nicht nur Platon namentlich, sondern führt auch zweimal den Werktitel an und gibt die entsprechende Passage in indirekter Rede wieder: ἐκεῖ γὰρ ὁ Σωκράτης φησί (denn dort sagt Sokrates). Expliziter und genauer hätte Aristoteles den intertextuellen Bezug kaum markieren können. Was hier so deutlich als Wiedergabe von fremdem Text ausgewiesen ist und also der erste Kandidat für eine Paraphrase (1), ist das Segment δεῖν κοινὰ τὰ τέκνα καὶ τὰς γυναῖκας εἶναι καὶ τὰς κτήσεις (dass die Nachkommen, die Frauen und die Besitztümer gemeinschaftlich sein müssen).

Allerdings ist dies nicht die einzige markierte Partie: ἐκεῖ (dort) nimmt ja die erste Nennung des Werkes durch ὥσπερ ἐν τῇ Πολιτείᾳ τῇ Πλάτωνος (wie in Platons *Politeia*) auf, was wiederum auf das vorhergehende καὶ τέκνων καὶ γυναικῶν καὶ κτημάτων κοινωνεῖν τοὺς πολίτας ἀλλήλοις (dass die Bürger untereinander Nachkommen, Frauen und Besitz gemeinschaftlich haben) zu beziehen ist. Für diesen Passus liegt also ebenfalls eine Markierung vor. Diese ist im Vergleich schwächer und legt, besonders wenn man die folgende Einleitung der indirekten Rede hinzunimmt, nahe, dass hier lediglich platonisches Gedankengut referiert wird und die eigentliche Paraphrase des platonischen Wortlautes erst noch folgt. Allerdings ist die Formulierung in großen Teilen derjenigen in (1) so ähnlich, dass auch für diese Passage (2) geprüft werden soll, inwiefern sie als Paraphrase betrachtet werden kann.

26 ἐπεὶ δὲ προαιρούμεθα θεωρῆσαι περὶ τῆς κοινωνίας τῆς πολιτικῆς, τίς κρατίστη πασῶν τοῖς δυναμένοις ζῆν ὅτι μάλιστα κατ' εὐχρὴν δεῖ καὶ τὰς ἄλλας ἐπισκέψασθαι πολιτείας, αἷς τε χρῶνται τινες τῶν πόλεων τῶν εὐνομεῖσθαι λεγομένων, κὰν εἴ τινες ἕτεροι τυγχάνουσιν ὑπὸ τινῶν εἰρημέναι καὶ δοκοῦσαι καλῶς ἔχειν [...]. Da es unser Vorhaben ist, Beobachtungen über die bürgerliche Gemeinschaft anzustellen, welche die beste von allen ist für diejenigen Menschen, die in der Lage sind, möglichst nach ihrem Wunsch zu leben, muss man auch die übrigen Verfassungen beobachten, die man teils in Staaten verwendet, von denen es heißt, sie seien in guter gesetzlicher Verfassung, teils in solchen, die von Denkern erörtert wurden und den Eindruck machen, in Ordnung zu sein [...]. (Übers. Schwarz)

Macht man sich in der *Politeia* auf die Suche nach dem Prätext, stößt man (als engste Parallele)²⁷ auf eine Stelle im fünften Buch, an der die vorher zunächst getrennt voneinander behandelten Bereiche der Gütergemeinschaft im Idealstaat, nämlich die familiäre und die materielle Gemeinschaft,²⁸ zusammengebracht werden (464 b5–c3):

τοῦ μεγίστου ἄρα ἀγαθοῦ τῇ πόλει αἰτία ἡμῖν πέφανται ἡ κοινωνία τοῖς ἐπικούροις τῶν τε παίδων καὶ τῶν γυναικῶν. [...] καὶ μὲν δὴ καὶ τοῖς πρόσθεν γε ὁμολογοῦμεν· ἔφαμεν γάρ που οὔτε οἰκίας τούτοις ἰδίας δεῖν εἶναι οὔτε γῆν οὔτε τι κτῆμα, [...] εἰ μέλλοιεν ὄντως φύλακες εἶναι.

So hat sich uns also die Gemeinschaft der Kinder und Frauen bei den Helfern als Ursache des größten Gutes für die Stadt erwiesen. [...] Und auch mit dem Früheren sind wir in Übereinstimmung: denn wir haben ja gesagt, dass diese weder eigene Häuser haben dürfen noch Land noch sonst einen Besitz, [...] wenn sie wirklich Wächter sein sollen. (Übers. nach Teuffel)

Im Vergleich des Passus mit den Formulierungen bei Aristoteles fällt zunächst auf, dass dieser die zwei Elemente der Gütergemeinschaft enger miteinander verbindet, nämlich die von Familie („die Nachkommen, die Frauen“) einerseits und die von materiellem Besitz („und die Besitztümer“) andererseits. Platon bringt diese wie gesagt erst hier zusammen, setzt sie aber noch durch die eingeschobene Zustimmung von Sokrates' Gesprächspartner Glaukon voneinander ab. Die relevanten Segmente bei Platon sind ἡ κοινωνία ... τῶν τε παίδων καὶ τῶν γυναικῶν (die Gemeinschaft ... der Kinder und Frauen) und ... (οὔτε οἰκίας) τούτοις ἰδίας δεῖν εἶναι ... οὔτε τι κτῆμα (dass diese [weder] ... haben dürfen ... noch sonst einen Besitz).

Vergleicht man nun die beiden Aristotelespartien miteinander, die im Verdacht stehen, Paraphrasen dieser Stelle zu sein, fällt auf, dass beide die platonische Wendung κοινωνία ... τῶν τε παίδων καὶ τῶν γυναικῶν (die Gemeinschaft der Kinder und Frauen), ergänzt durch τι κτῆμα (einen Besitz) aus dem 2. Absatz, in

27 Platon behandelt das Thema an mehreren Stellen, doch ist keine hinreichend ähnlich, um sie eindeutig als Vorlage der Formulierung bei Aristoteles identifizieren zu können. Möglicherweise suggeriert er durch die Markierung lediglich, dass er sich auf einen konkreten Passus bezieht, fasst aber eigentlich allgemeiner die von Sokrates vertretene Position zusammen oder paraphrasiert aus dem Gedächtnis. Natürlich fiel aber auch ein solches Vorgehen unter das Phänomen des Paraphrasierens.

28 Die Vorschrift zur Gemeinschaft der Wächter in Bezug auf Frauen und Kinder wird im fünften Buch ab 457 c diskutiert und mündet in dieser Zusammenfassung mit der Gütergemeinschaft, welche bereits im dritten Buch (bis 417 b) behandelt ist.

recht enger sprachlicher Anlehnung wiedergeben. Hierbei entspricht die Genitivkonstruktion in (2) eher der platonischen Syntax, dafür sind in (1) die Artikel und die Wortstellung beibehalten. Davon abgesehen ist an den Aristotelesstellen der Fokus leicht unterschiedlich gesetzt: Während in (1) durch $\delta\epsilon\acute{\iota}\nu\ \epsilon\acute{\iota}\nu\alpha\iota$ das Element der Notwendigkeit in Platons Formulierung aus dem zweiten, auf die Gütergemeinschaft bezogenen Satz wörtlich aufgenommen ist, findet sich in (2) eine Information, die in (1) ausgelassen ist, nämlich für wen die Gemeinschaft gelten soll. Hier hat Aristoteles allerdings die platonische Konzeption abwandelt (wenn nicht verfälscht): Statt nur auf die Wächter, von Platon hier als ‚Helfer‘ bezeichnet ($\tau\omicron\iota\varsigma\ \epsilon\pi\iota\kappa\omicron\upsilon\tau\omicron\iota\varsigma$), wendet er das Konzept auf $\tau\omicron\upsilon\varsigma\ \pi\omicron\lambda\iota\tau\alpha\varsigma$, also alle Bürger, an (betont durch das hinzugefügte $\acute{\alpha}\lambda\lambda\eta\lambda\omicron\iota\varsigma$ ‚untereinander‘).

Unterstützt durch die Visualisierung in [Abb. 1](#) und [2](#) kann man so kurz und unsystematisch das Verhältnis der Partien (1) und (2) zum Prätext beschreiben und sie unter diesem Aspekt miteinander vergleichen. Aufgrund dieser Beschreibung und auch wenn (1) viel deutlicher als Paraphrase gekennzeichnet ist, fällt es schwer zu entscheiden, welche der Passagen näher am platonischen Prätext liegt. Bei dem Versuch, einen solchen Paraphrasen-Vergleich zu systematisieren und auf eine theoretische Grundlage zu stellen, zeigen sich schnell Probleme:

Offensichtlich sind die allgemeineren Definitionen des Paraphrasenbegriffs nach dem Kriterium der semantischen Übereinstimmung bzw. Nähe hier wenig hilfreich. Im Vergleich mit der Platonstelle liegt z.B. „approximate conceptual equivalence“²⁹ bei Aristoteles sowohl in (1) als auch in (2) vor. Um aber bestimmen zu können, an welcher der beiden Stellen die Annäherung an die „conceptual equivalence“ größer ist, müsste man nicht nur benennen können, worin diese besteht (also wissen, was als Konzept zugrunde liegt), sondern sie auch bewerten, also messen können. Auch die Identifizierung und Beschreibung der verschiedenen Transformationen vom Prä- zum Posttext bringt höchstens eine Annäherung, wirft aber zugleich Probleme auf. So ist in unserem Fall sicherlich strittig, ob man $\pi\omicron\lambda\iota\tau\alpha\varsigma$ (die Bürger) als synonym zu $\epsilon\pi\iota\kappa\omicron\upsilon\tau\omicron\iota\varsigma$ (die Helfer) ansehen darf. Daran, dass die meisten das aus gutem Grund verneinen würden, werden gleich zwei Probleme dieses Ansatzes deutlich: dass solche semantischen Kategorien oft zu enge Schubladen sind und dass sie auf subjektiver Beurteilung fußen.³⁰

29 Vgl. oben [S. 35](#), [Anm. 27](#).

30 Das lässt sich auch an Platons eigener Formulierung veranschaulichen: Innerhalb des betroffenen Passus und an anderen Stellen benutzt er alternativ die Bezeichnung ‚Helfer‘ für die Wächter seines Idealstaates; innerhalb dieses fiktiven Rahmens sind die Begriffe also synonym gesetzt, obwohl sie das im normalen Sprachgebrauch nicht sind. Alternativ könnte man die Ersetzung der Helfer durch die Bürger in der Aristotelespassage (2) noch als Transformation auf Grundlage einer Teil-Ganzes-Relation interpretieren. Das wäre weniger problematisch, aber lässt dennoch keine Schlüsse über die semantische Nähe zu. Bei der Distanzmessung wies diese Relation mit 1.13 den höchsten Wert (also die geringste Ähnlichkeit) auf, vgl. [Tabelle 1](#) unten.

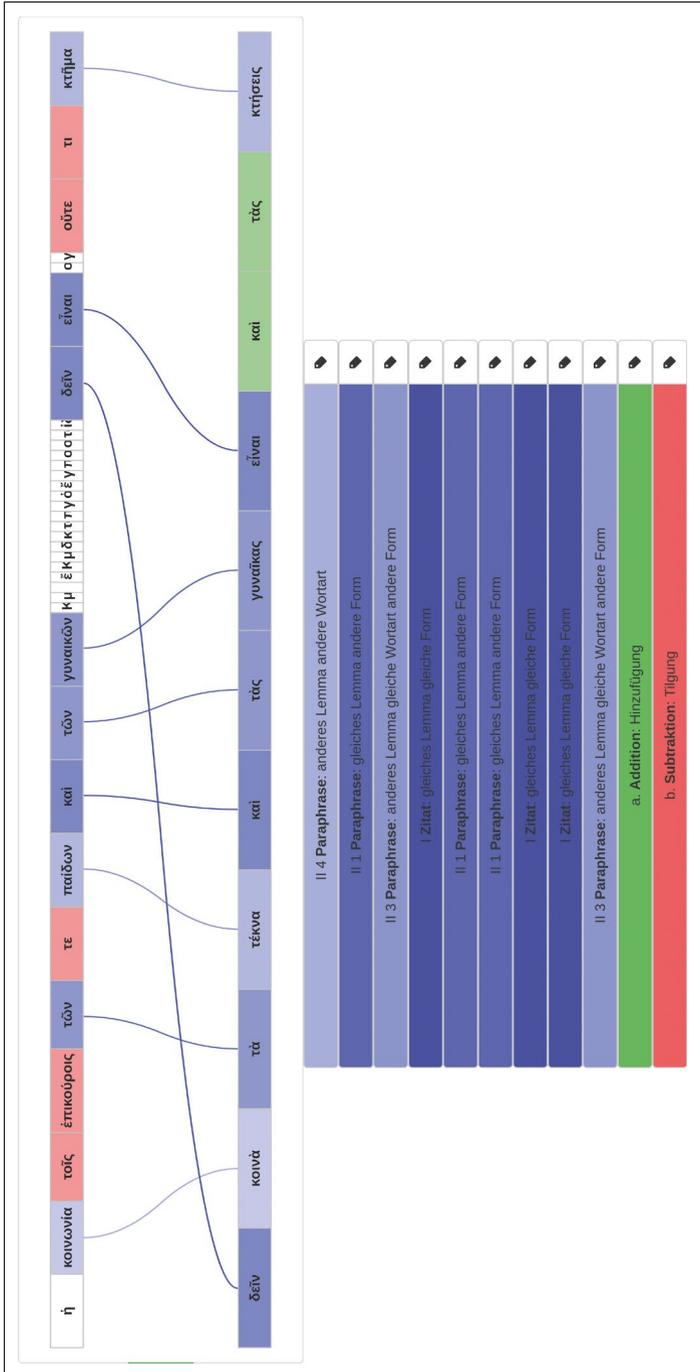


Abbildung 1. Visualisierung von Wort-zu-Wort-Relationen sowie von Additionen (grün)/Subtraktionen (rot) zwischen der *Politeia*-Stelle (oben) und dem Aristotelespassus (I) mithilfe des Referenzannotierers (vgl. zum Tool Pöckelmann / Wöckener-Gade [2018]). Es wurden 9 Relationen (hier blau) annotiert: 3 × Zitatanteil, 6 × Paraphrasenanteil; vier dieser Relationen bestehen zwischen Stoppwörtern.

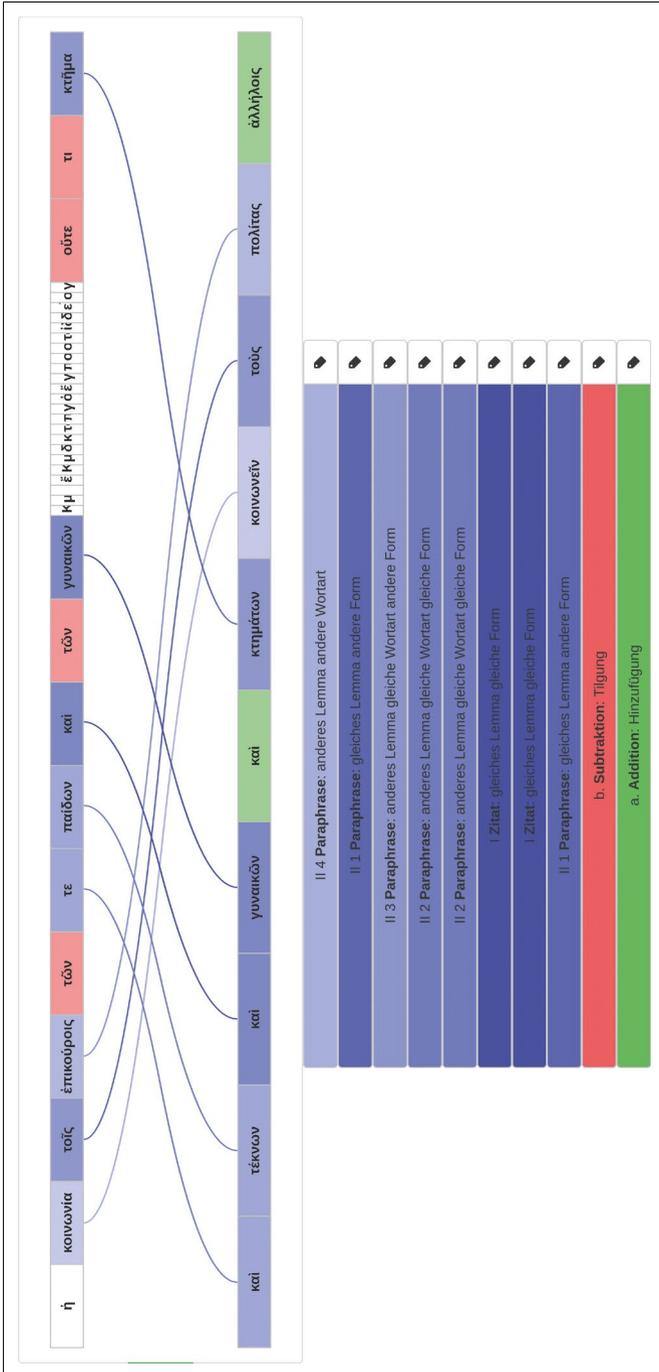


Abbildung 2. Visualisierung von Wort-zu-Wort-Relationen sowie von Additionen / Subtraktionen zwischen der *Politeia*-Stelle (oben) und dem Aristotelespassus (2) mithilfe des Referenzannotierers. Es wurden acht Relationen annotiert: 2 × Zitatanteil (wortwörtliche Übereinstimmung), 6 × Paraphrasenanteil; drei dieser Relationen bestehen zwischen Stopppwörtern.

Zudem bleibt eine solche Kategorisierung selbst bei recht eindeutig zu bestimmenden Transformationen³¹ rein deskriptiv. Die Frage, ob zu Platons *κοινωνία* (die Gemeinschaft) in unserem Fall das adjektivische *κοινά* in (1) oder das verbale *κοινωνεῖν* in (2) die engere Paraphrase darstellt, ist so nicht zu beantworten. Auch könnte man höchstens additiv die Zahl der Transformationen anführen, woraus sich aber wiederum kaum Schlüsse über deren Nähe zum Prätext ziehen ließen.

Für den Charakter der Paraphrase scheint vielmehr die semantische Ähnlichkeit zum Prätext konstituierend,³² die bisher mit traditionellen Methoden allerdings schwer bestimmbar erschien. Als Messverfahren, wie es sich Pfister für die Intensität von Intertextualität wünschte, bieten sich für die Ähnlichkeitsbeziehung der Paraphrase selbst zu ihrem Prätext nun die innerhalb der Digital Humanities inzwischen etablierten Distanzmaße an. Besonders geeignet erscheint hierfür das Vektorraum-basierte Verfahren der Word Mover's Distance (WMD), das mit wenigen Vorannahmen auskommt, um semantische Ähnlichkeit abzubilden.³³ Um für unser Beispiel die Frage zu beantworten, welche der Aristotelesstellen die nähere Paraphrase darstellt, wurde die Distanz beider zu der Platonstelle gemessen. Der Vergleich beschränkt sich hierbei auf die relevanten Begriffspaare (jeweils fünf). Das hängt teilweise mit der Funktionsweise des Verfahrens zusammen: Da die WMD prinzipiell Wort zu Wort zuordnet, erhöht z. B. ein Begriff im Prätext, dem im Posttext nichts zugeordnet werden kann, die Distanz unverhältnismäßig. Stoppwörter wurden als semantisch unbedeutend nicht berücksichtigt, aus eben diesem Grund können sie auch vom Distanzmaß nicht zuverlässig zugeordnet werden.³⁴

Die Abstandsmessung ergab – wie zu erwarten – vergleichsweise niedrige Werte sowohl für den Abstand von (1) als auch (2) vom platonischen Prätext. Die beiden Stellen sind dem Prätext wie auch sich untereinander in der Tat sehr ähnlich. Den geringeren Abstand zur Vorlage weist in dieser Messung aber die Variante (2)

31 Als Beispiel für solche können die in den Visualisierungen annotierten Relationen dienen: Um Subjektivität soweit wie möglich auszuschließen, werden die Transformationen mit formalen, lexikalisch-morphologischen Kategorien beschrieben: Als Zitatanteil gilt dabei die Übereinstimmung von Lemma und Form, Paraphrasen reichen von der Übereinstimmung des Lemmas, aber Abweichung in der Form bis zu der Ersetzung eines Lemmas durch mehrere andere bzw. umgekehrt.

32 Vgl. auch Vila et al. (2014) 209 f., die zunächst mit Verweis auf die Forschungslage die Möglichkeit von kompletter semantischer Übereinstimmung („sameness“) zwischen Paraphrase und Prätext kritisch hinterfragen, um dann zu folgern: „Therefore, paraphrasing must be situated in the field of the approximation, opening the path to different semantic similarity degrees or degrees of paraphrasability. Paraphrasing takes place in a continuum that goes from absolute identity to the absence of semantic similarity.“

33 Vgl. zum Verfahren für das Altgriechische in diesem Band den [Beitrag von Molitor/Ritter und Pöckelmann S. 45–60](#).

34 Vgl. hierzu den [Beitrag von Rautenberg in diesem Band S. 111–123](#).

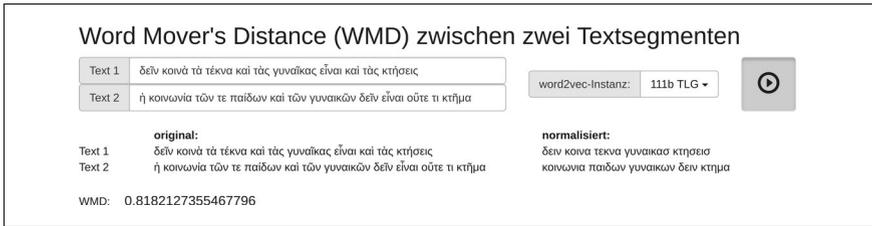


Abbildung 3. Distanzmessung via WMD zwischen der *Politeia*-Stelle (oben) und dem Aristotelespassus (1)

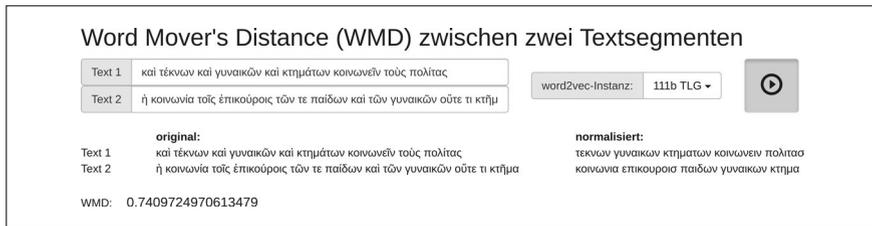


Abbildung 4. Distanzmessung via WMD zwischen der *Politeia*-Stelle (oben) und dem Aristotelespassus (2)

auf, also die nicht eindeutig als Bezugnahme auf den platonischen Wortlaut gekennzeichnete Partie (0.74 im Vergleich zu 0.81), s. Abbildungen 3 und 4.³⁵ Es lassen sich auch Distanzen einzelner Wortpaare errechnen (s. [Tabelle 1](#)) und so die Transformationen bewerten.

Beispielsweise ist die Distanz zwischen *κοινωνία* (Gemeinschaft) bei Platon und *κοινά* (gemeinschaftlich) in (1) mit 1.07 höher als diejenige zu *κοινωνεῖν* (gemeinschaftlich haben) in (2) mit 0.94; die Paraphrase mittels des Verbalausdrucks liegt also näher an der Vorlage. Interessant mit Blick auf das Verfahren ist auch, dass die geringste Distanz einer abweichenden Formulierung mit 0.68 diejenige zwischen *παίδων* (der Kinder) bei Platon und *τέκνων* (der Nachkommen) in (2) ist. Die Ersetzung durch ein anderes, synonymes Lemma in derselben Form ist hier als engere Entsprechung gemessen als die Beibehaltung des Lemmas und Änderung der Form in zwei anderen Fällen (*γυναικῶν* → *γυναῖκας* (1): 0.8; *κτῆμα* → *κτημάτων* (2): 0.95), was darauf hindeutet, dass sich formale lexikalische und syntaktische Kriterien nicht zwangsläufig mit der gemessenen semantischen Nähe korrelieren lassen.

35 Die Word Mover's Distance beruht mathematisch auf der euklidischen Distanz. Die Werte sind zudem von der Implementierung des Korpus mittels Word2Vec und weiteren Parametern des Algorithmus abhängig. Da noch nicht genügend Vergleichsdaten zur genaueren Einordnung vorliegen, mag zum Vergleich die Distanz zwischen den semantisch wenig nahen Begriffen *λάμπειν* (leuchten) und *πόλεμος* (Krieg) dienen, die mit derselben Methode gemessen 1.41 betrug.

Tabelle 1. Die paraphrasierten Wörter des Prätextes und die entsprechenden Reformulierungen in (1) und (2) mit Angabe der Word Mover's Distance³⁶

Platon Wort	Aristoteles (1) Wort: WMD	Aristoteles (2) Wort/ WMD
κοινωνία	κοινά: 1.07	κοινωνεῖν / 0.94
ἐπικούροις	-	πολίτας / 1.13
παίδων	τέκνα: 1.06	τέκνων / 0.68
γυναικῶν	γυναῖκας: 0.8	γυναικῶν / 0
δεῖν	δεῖν: 0	-
κτῆμα	κτήσεις: 1.16	κτημάτων / 0.95
WMD der gesamten Passage (ohne Stoppwörter)	0.81	0.74

Laut dem Ergebnis der Distanzmessung steht Passage (2) zumindest in den Schlüsselbegriffen Platons Text semantisch näher und kann in diesem Sinne um nichts weniger (sondern sogar ein bisschen mehr) als Paraphrase der platonischen Formulierung gedeutet werden, selbst wenn Aristoteles das nicht beabsichtigt haben sollte, wie seine Formulierung nahe legt, da er ausdrücklich nur die folgende Paraphrase (1) als Bezugnahme auf den Wortlaut auszeichnet.³⁷ Bezieht man allerdings formale Kriterien wie die Wortstellung oder die Setzung der Artikel mit ein, erscheint in diesen Punkten (1) enger an den Prätext anzuschließen. Gerade aufgrund dieser Diskrepanz verdeutlicht das Beispiel das Potenzial des Verfahrens bei der Untersuchung und Charakterisierung von Paraphrasen und verwandten Phänomenen.

Auf Basis der Ähnlichkeitsmessung können Paraphrasen nicht nur wie im Beispiel miteinander verglichen werden, auch ihr Standort innerhalb des Phänomens ‚Paraphrase‘ ist so genauer zu bestimmen. Hierbei dient auf der einen Seite das Zitat mit der kleinsten möglichen Distanz vom Prätext als Grenze,³⁸ auf der anderen stehen Formulierungen, die sich nicht mehr auf ein konkretes Textsegment

36 WMD-Paarvergleich mit der Word2Vec-Instanz 111b (ohne Stoppwörter, Diakritika und Groß-/Kleinschreibung normalisiert), abgerufen am 26.10.2018. Die Werte sind auf zwei Nachkommastellen gerundet.

37 Hieran zeigt sich einmal mehr, dass die Intention des Autors bei der Bewertung von intertextuellen Bezügen ein mindestens unsicherer Wegweiser ist.

38 In einer Distanzmessung entspräche diese logischerweise dem Wert 0; das Verfahren deutet darauf hin, dass nur in dem Fall, dass Passagen im Wortlaut miteinander identisch sind, man auch von völliger semantischer Übereinstimmung ausgehen darf.

sondern beispielsweise zusammenfassend auf ein Gesamtwerk beziehen, so dass sprachliche Ähnlichkeit zwischen Prä- und Posttext aufgrund der formalen Unterschiede schlicht nicht detektierbar sind. Während man solche Erscheinungen mit einem unscharfen, sehr weiten Paraphrasenbegriff zwar noch einem Grenzbereich des Phänomens zuordnen kann (selbst wenn derart komplexe Transformationen sich der Analyse entziehen), ist der Bereich der Nachwirkung, die, soweit möglich, die Inhalte losgelöst von der sprachlichen Form zu transportieren sucht, von dem, was Paraphrase umfasst, abzugrenzen. Zwischen diesen beiden Polen, dem Zitat und der Nachwirkung, spannt sich demnach das Feld auf, innerhalb welchem Paraphrasen einzuordnen sind. Hierbei wäre grundsätzlich zwischen Paraphrasen, die sich nah an der Wortebene ihres Prätextes befinden (enge Paraphrase), und solchen, die sich weiter von dieser entfernen, aber noch signifikante Ähnlichkeiten ausweisen, zu unterscheiden. Im Vergleich zu früheren Ansätzen scheint der Versuch einer Charakterisierung unter Zuhilfenahme der mathematischen Abstraktionsebene und anhand einer Ähnlichkeitsbestimmung, wie sie das Messverfahren der Word Mover's Distance ermöglicht, einige Vorteile zu bieten, die die neue Methode zumindest als sinnvolle Ergänzung der traditionellen Forschungspraxis erscheinen lassen: Denn was in den Texten selbst an intertextuellen Spuren liegt, sind neben Signalen im Kontext wie Markierungen eben relevante Ähnlichkeiten von Post- und Prätext, und diese können mithilfe des Verfahrens reproduzierbar sowohl gefunden als auch gemessen werden. Durch diese Messbarkeit wird Ähnlichkeit also zu einer skalierbaren Größe, die eine Abstufung von Paraphrasen und anderen intertextuellen Phänomenen unter dem Aspekt des Mehr oder Weniger erlaubt. Die informatischen Hilfsmittel können daher nicht nur eingesetzt werden, um Textähnlichkeiten in einer objektiven, operationalisierbaren Weise zu finden, sondern auch zu deren Analyse herangezogen werden. Sie ermöglichen es dabei, die Kategorisierung einer Paraphrase anhand der Nähe zum Prätext als ‚enge‘ oder ‚weite‘ Paraphrase mit einem Zahlenwert zu verbinden. Dieser Befund wäre in einem zweiten Schritt mithilfe weiterer Methoden einzuordnen und zu interpretieren, um das Verfahren heuristisch fruchtbar zu machen. So kann man aus der linguistischen Perspektive innerhalb der Paraphrase einzelne Transformationstypen nach formalen oder semantischen Kriterien analysieren, wie sie u. a. von Vila et al. vorgeschlagen wurden. Innerhalb einer eher literaturwissenschaftlich orientierten Fragestellung kann man die Paraphrase im Kontext literarischer Techniken wie Anspielung oder Plagiat beschreiben oder in Anlehnung an Pfister den intertextuellen Charakter durch eine Analyse der kontextuellen Einbettung näher bestimmen.

Abschließend sei nochmals betont, dass sich der eigentliche intertextuelle Bezug immer erst dadurch, dass er vom Rezipienten hergestellt wird, konstituiert, weshalb ein subjektives Moment, wenn man nicht in einem kruden Positivismus

enden will, in Intertextualitätsfragen und deren Erforschung unumgänglich ist. Was in diesem Beitrag vorgeschlagen wird, ist vor allem dies: Die traditionelle geisteswissenschaftliche Forschung wird durch die informatischen Methoden keineswegs entwertet, sondern gewinnt mit ihrer Einbeziehung vielmehr eine neue Erkenntnismöglichkeit hinzu, die sie der Objektivitätsforderung an die Wissenschaft ein wenig näherbringt.

