

DHLehre

Integration von Forschung und universitärem Unterricht

Thomas Krefeld

Forschungskommunikation

Die universitäre Lehre folgt – vom speziellen Fall der Didaktik abgesehen – der Forschung (und nicht etwa umgekehrt). Mit der Frage, was unter Forschung zu verstehen sei, brauchen wir uns nicht lange aufzuhalten. Einen guten Hinweis gibt bereits die Herkunft der englischen Bezeichnung *research*, die – wie jedes gute englische Wort – aus dem Fra. stammt, nämlich aus *recherche*, das wiederum vom Verb *rechercher* mit der ursprünglichen Bedeutung ‚wiederholt, dauerhaft suchen‘ abgeleitet wurde. In diesem Sinn lässt sich Forschung immer noch als Suche nach fehlenden, bzw. adäquaten Wissensbeständen umschreiben: Forschung sucht methodisch nachvollziehbare Antworten auf theoretisch begründete Fragen.

Der Drang nach Wissenvermehrung, d.h. nach Forschung ist zweifellos eine anthropologische Konstante, wie Michel de Montaigne (Bd. III, Kap. 13, *De l'expérience*) in Fortführung eines aristotelischen Prinzips¹ festgestellt hat:

„Il n'est desir plus naturel que le desir de cognoissance.
(„Es gibt kein natürlicheres Verlangen als das nach Wissen.“)

Alles andere als konstant sind jedoch die Rahmenbedingungen der Forschung, denn sie sind massivem historischen Wandel unterworfen; dazu gehören die Institutionen, die Finanzierungswege und nicht

1 <<All men naturally desire knowledge.>> (Aristotle 1933, 980a22)

zuletzt die Formen der Kommunikation. Vor allem letztere wurden seit dem Aufkommen der Neuen Medien revolutioniert: Die Gutenberg-Galaxis (vgl. McLuhan 1962) strebt mit zunehmender, rasanter Geschwindigkeit auseinander, da die Gravitationskraft ihres Fixsterns, des gedruckten Buchs, immer schwächer wird; auch von der ehemaligen Leuchtkraft ist nurmehr ein matter Schimmer geblieben. In Forschungsprojekten, die mit Webtechnologie operieren, verläuft die Wissenschaftskommunikation jedenfalls nach vollkommen anderen Regeln (vgl. Krefeld 2017c), die hier nicht im Einzelnen ausbuchstabiert werden müssen. Es reicht im Hinblick auf die Lehre auf die vollkommen veränderte Realität der Publikation hinzuweisen. Die traditionelle Kommunikation von Forschungsergebnissen lässt sich wie folgt schematisieren:

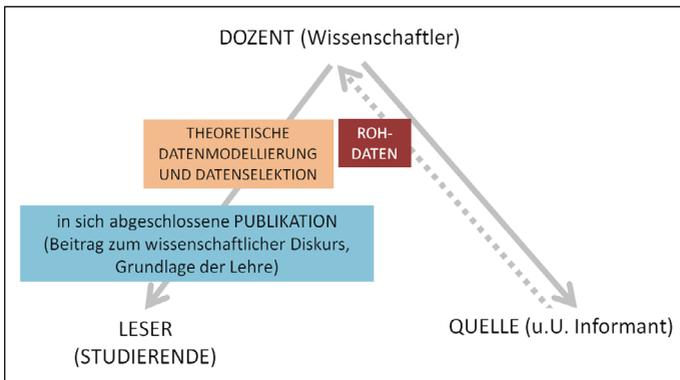


Abbildung 1: Forschungskommunikation auf der Grundlage gedruckter Texte

Publikationen, die für den Druck konzipiert werden, sind gewissermaßen dingliche, objektive Endprodukte eines Forschungsprozesses. Ihre (kommerzielle) Herstellung und Distribution liegt in der Hand Dritter, der Verlage, und ihre Form entspricht spezifischen Gattungen, wie zum Beispiel einem Wörterbuch, einer Abhandlung, einer Anthologie, einem Atlas oder auch einem Lehrbuch.

Unter den Bedingungen des Web lässt sich der wissenschaftliche Diskurs im Rahmen von freier Zugänglichkeit (*open access*) und offenen Quellcodes (*open source*) etwa wie folgt schematisieren:

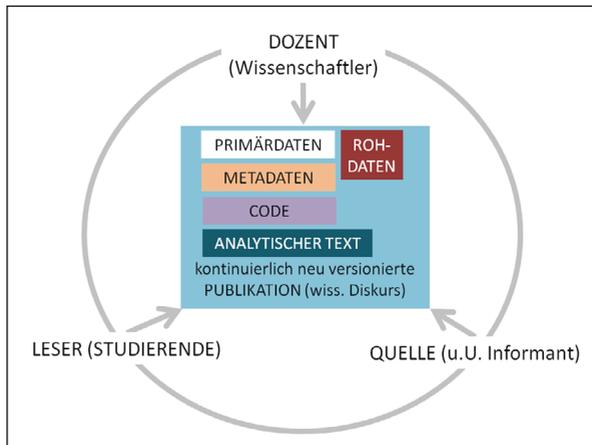


Abbildung 2: Webbasierte Forschungskommunikation

Unter diesen Bedingungen erweitert sich zunächst der Bereich des Publizierten, da außer dem analytischen Fließtext auch die Primär- und Metadaten also solche sowie speziell entwickelter Code veröffentlicht werden. Sodann sind die Gattungsgrenzen aufgehoben, allerdings ohne die sinnvolle Funktionalität der traditionellen Gattungen zu entwerfen. So sind im Fall des DFG-Projekts *Verba Alpina* zwar ein Wörterbuch (*Lexicon Alpinum*), eine Atlasdarstellung (*Interaktive Karte*) und analytischer wissenschaftlicher Text (*Methodologie*) als eigenständige Module angelegt; untereinander sind sie jedoch so verlinkt, dass von einer Funktion stets in eine andere gewechselt werden kann.²

Die teils beharrlichen Vorbehalte gegenüber Publikationen im Web werden durch ganz unterschiedliche Gründe getragen, die allesamt nicht substantiell, sondern akzidentiell zu sehen sind. Am schwersten wiegt das Argument der prekären Nachhaltigkeit (vgl. dazu Krefeld &

2 https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=162

Lücke 2017), das durch ein Paket konzertierter Maßnahmen bewältigt werden kann:

- die Schaffung textstabiler und insofern dauerhaft zitierfähiger Versionen;
- die Auszeichnung mit Metadaten, die sowohl durch Suchmaschinen wie in Bibliothekskatalogen leicht auffindbar sind;
- die Ablage auf institutionell betriebenen Servern mit zuverlässiger Aussicht auf langfristige Verfügbarkeit (große Rechenzentren).

Ein Format, das diesen Ansprüchen genügen kann, wurde an der LMU in Zusammenarbeit mit der UB mittlerweile entwickelt. Es ist so angelegt, dass die jeweilige textstabile Version in der URL angezeigt wird und weiterhin durch die Nummer eines Abschnitts, auf den man sich bezieht, oder den man wörtlich zitiert, spezifiziert werden kann. Hier die empfohlene Zitierweise eines Artikels in der 7. Version (...v=7):

- Thomas Krefeld & Stephan Lücke (2017): Nachhaltigkeit – aus der Sicht virtueller Forschungsumgebungen. Korpus im Text. Version 7 (10.03.2017, 12:27). url: <http://www.kit.gwi.uni-muenchen.de/?p=5773&v=7>

sowie die URL seines ersten Abschnitts (...p:1):

- <http://www.kit.gwi.uni-muenchen.de/?p=5773&v=7 - p:1>

Der soeben zitierte Text wurde in unserer Publikationsplattform *Korpus im Text* (KiT) veröffentlicht, die nicht zuletzt für die Lehre relevant ist. Ein Lehrbuch, in das vielfältige Forschungserfahrungen eingeflossen sind, sowie mehrere wissenschaftliche Arbeiten und Artikel stehen hier bereits zur Verfügung; ein weiteres Lehrbuch, einige Dissertationen und andere, auch umfangreiche wissenschaftliche Vorhaben werden vorbereitet.³ Darüber hinaus bietet auch DH-Lehre spezi-

3 http://www.kit.gwi.uni-muenchen.de/?post_type=band

fische Publikationsmöglichkeiten mit engem Bezug zur Lehre, nämlich für exzellente [Abschlussarbeiten](#) einerseits und für ausformulierte [Vorlesungen](#) andererseits.

Ein neues Kooperationsprinzip

Im Hinblick auf die Lehre ist jedoch vor allem von Bedeutung, dass sich mit der webbasierten Forschungs- und Wissenschaftskommunikation neue Horizonte der akademischen Zusammenarbeit von lehrenden Forschern und Lernenden auftun. Ganz selbstverständlich ergeben sich Kooperationen zwischen Dozenten, bzw. Forschungsprojekten und Studierenden einerseits und zwischen Studierenden untereinander andererseits. Dazu drei Beispiele.

1. Kooperation zwischen Projekten und Studierenden

Im Rahmen einer Master-Arbeit wurden von Myriam Abenthum lexikologische Kommentare verfasst, die nach Revision durch die Projektleiter in das [Wörterbuchmodul](#) des Projekts *Verba Alpina* übernommen wurden.

2. Kooperation zwischen Doktoranden und BA-Studierenden zur Toolentwicklung

Etwa ein Drittel der sprachwissenschaftlichen DoktorandInnen unserer Graduiertenschule nimmt die Lehre der ITG in Anspruch und stützt ihre Dissertation auf relationale Datenbanken;⁴ die Themenbereiche sind sehr unterschiedlich und die BetreuerInnen weit gestreut. Darunter ist ein ebenso innovatives wie methodisch richtungsweisendes Projekt von Sebastian Lasch zur Erfassung und Analyse von Graffiti im Stadtgebiet Roms. Bei der kartographischen Visualisierung der ca. 3000 Graffiti werden Instrumente zur Darstellung quantitativer Verhältnisse zum Einsatz kommen, die für *Verba Alpina* entwickelt wurden (insbesondere eine *heatmap*). Aber eine nicht geringe Schwierigkeit bei der Datenerfassung und -verarbeitung besteht darin, die unterschiedlichen Information so aufzunehmen, dass jedes Zeichen

4 <https://www.dh-lehre.gwi.uni-muenchen.de/?p=57328>

einzelnen mit spezifischen Metadaten und darüber hinaus auch im Kontext der Gesamtäußerungen und ihrer Abfolge wiedergegeben werden kann. Das folgende Foto zeigt:

- von einer Hand den grünen Text *Stalin c'è* („Stalin ist da“) mit zwei grünen Hammer & Sichel-Symbolen;
- von einer zweiten Hand zwei schwarze Zusätze, die den grünen Text negieren: *Stalin non c'è più* („Stalin ist nicht mehr da“);
- eventuell von einer weiteren Hand (?) die Übermalung der Hammer & Sichel-Symbole durch faschistische Kreis & Kreuz-Symbole;
- die Übermalung von *non* und *più* mit dem offenkundigen Ziel sie unkenntlich zu machen und die Negation aufzuheben;

dem Anschein nach zuletzt angebrachte weiße Übermalungen, ein weißes faschistisches Symbol über dem schwarz übermalten, ursprünglich grünen sowie eine unleserliche weiße Schrift über dem Namen Stalin.



Abbildung 3: Eine Wand mit Graffiti in Rom (Foto Sebastian Lasch)

Eine direkte Aufnahme dieser Daten in das relationale Modell, mit dem die ITG ausnahmslos arbeitet erschien möglich, jedoch reichlich umständlich. Just als diese Überlegungen angestellt wurden, tauchte ein Studierender der Informatik an unserem Institut (der Romanistik) auf, Daniel Pollithy, und bekundete sein Interesse, in seiner informatischen BA-Arbeit etwas zu entwickeln, das sich auf sprachliche Daten anwenden ließe. Ich skizzierte ihm das Problem und in Abstimmung mit der ITG und engem Austausch mit dem Doktoranden entstand ein außerordentlich nützliches Tool, das eine direkte Erfassung der Primärdaten und die Zuordnung von Metadaten in beliebige vielen Layern und ausgehend vom Foto selbst gestattet. Einen Eindruck der Bearbeitung (aus der Alpha-Version) gibt der folgende Screenshot:



Abbildung 4: Erfassung semiotischer Informationen auf Basis einer Photographie (Sebastian Lasch | Daniel Pollithy)

Das Datenschema zeigt, dass bei der Erfassung die syntaktische Abfolge (FOLGT) unterschiedlicher Zeichentypen (TOKEN, SYMBOL), die zeitliche Abfolge (ÜBERDECKT), die Assoziation mit semantischen bzw. kognitiven Feldern (AKTIVIERT) sowie diverse semiotische und linguistische Metadaten berücksichtigt werden.

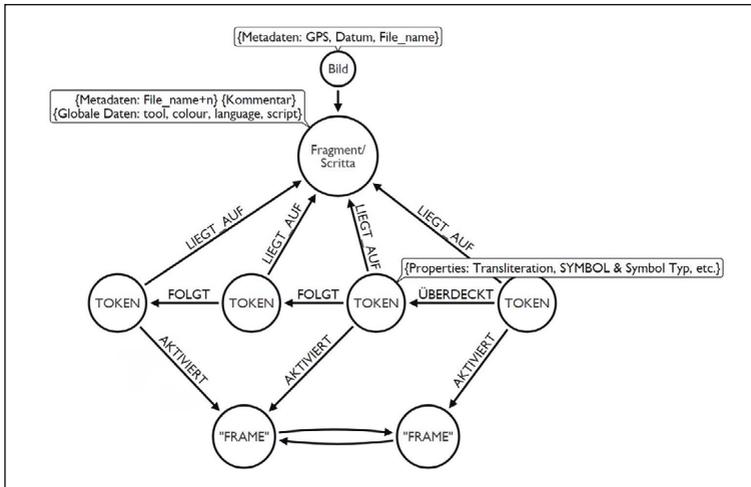


Abbildung 5: Datenmodell für die Erfassung von Graffiti
(Sebastian Lasch | Daniel Pollithy)

Eine syntaktische spezifizierte Verknüpfung („follows“) zeigt der folgende Screenshot:

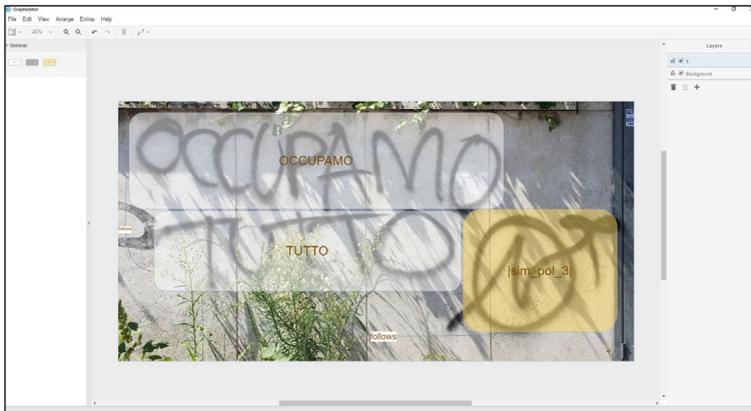


Abbildung 6: Syntaktische Relation zwischen semiotischen Einheiten
(Sebastian Lasch | Daniel Pollithy)

Für das präsentierte Tool wurde zwar ein *graph database* verwandt (die Community-Edition der Neo4j GraphDB); es gestattet jedoch den Export der erfassten Daten in das übliche relationale Modell (MySQL) der ITG. Nach Fertigstellung der BA-Arbeit wird es dem DHVLab zur Verfügung gestellt und sich dort sicherlich als sehr nützlich für editorische, lexigraphische und andere Projekte erweisen.

Kooperation von Forschung Lehre im Bereich der Korpora und Repositorien

Ein letztes Beispiel illustriert eine gleichermaßen strategische wie strukturelle Form der Kooperation unserer Lehr- und Forschungsinfrastruktur. Sie besteht darin, Korpora, die für Dissertationsprojekte konzipiert, erhoben und strukturiert werden, nach Abschluss der Dissertation der Lehre (und natürlich weitere Forschung) zur Verfügung zu stellen. Ein prototypisches Beispiel ist das italienische WhatsApp-Korpus, das Katharina Jakob im Rahmen einer Untersuchung über *Medienbedingte Variation am Beispiel italienischer WhatsApp-Chats* aufgestellt und mit einem soliden Datenbestand gefüllt hat, wie der folgende Screenshot zeigt:

Zeige Datensätze 0 - 99 (43583 insgesamt, Die Abfrage dauerte 0.0012 Sekunden.)

SELECT * FROM `whatsapp_chats`

1 > >> Anzahl der Datensätze: 100 Zeilen filtern: Diese Tabelle durchsuchen Nach Schlüssel sortieren

+ Optionen							
	msg_id	ora	data	alias	testo	Quelle	wordcount
<input type="checkbox"/>	1	12:08	25/set/2014	alessandra	La iellatura continua	chat001	3
<input type="checkbox"/>	2	12:08	25/set/2014	alessandra	Oggi esco di casa e uscendo mi porto l'immondizia...	chat001	29
<input type="checkbox"/>	3	12:08	25/set/2014	alessandra		chat001	0
<input type="checkbox"/>	4	12:40	25/set/2014	Zoe	noooo	chat001	1
<input type="checkbox"/>	5	12:40	25/set/2014	Zoe	<Media omesso>	chat001	2
<input type="checkbox"/>	6	12:44	25/set/2014	alessandra	Una lingua sconosciuta	chat001	3
<input type="checkbox"/>	7	12:45	25/set/2014	Zoe	<Media omesso>	chat001	2
<input type="checkbox"/>	8	12:46	25/set/2014	Zoe	passagai significa una che passa i guail	chat001	7
<input type="checkbox"/>	9	12:46	25/set/2014	alessandra	Sono	chat001	4

Abbildung 7: Allgemeine Ansicht der Datensätze im WhatsApp-Korpus von Katharina Jakob (Screenshot)

Diese spezifische Art des aktuellen Sprachgebrauchs zeichnet sich durch etliche, erst ansatzweise analysierte Besonderheiten aus, wie z.B. durch das systematische Fehlen des Apostrophs nach der Kurzform des definiten Artikels vor vokalischem beginnendem Folgewort: *il*, *la* > *l* anstatt der Standardform *l'*; das folgende Bild zeigt die Abfrage und den Beginn der Trefferliste.

The screenshot shows the phpMyAdmin interface with a SQL query and its results. The query is:

```
SELECT * FROM `whatsapp_chats` WHERE testo REGEXP '(?![a-zA-Z])' OR testo REGEXP '(?![a-zA-Z])'
```

The results table is as follows:

msg_id	ora	data	alias	testo	Quelle	wordcount
3897	06:55	01/dic/2014	Diletta	Bella l idea	chat006	3
3944	20:45	03/dic/2014	Diletta	...ora mi sto scaldando spinaci l unica cosa che ...	chat006	15
4269	23:43	11/dic/2014	Diletta	Domani sera pensavo di fare un bar tour con la [no...	chat006	58
4612	18:42	10/dic/2014	Diletta	Mi sn confusa con l altro tu collega	chat006	8
4653	21:19	10/dic/2014	Diletta	Come va l aperitivo? Ho fatto un insalata e una b...	chat006	17
4693	19:31	11/dic/2014	Diletta	L autista potrebbe essere ubriaco. Guida come una...	chat006	12
4752	20:38	12/dic/2014	Diletta	Si paga l entrata al [luogo] stase quindi?	chat006	8
4841	17:41	13/dic/2014	Diletta	Siamo nel duomo ho trovato per terra dove esce l a...	chat006	22
5134	14:24	16/dic/2014	Diletta	E stato l unico tramezzino che c era...mensa chiu...	chat006	11
5258	02:29	21/dic/2014	Diletta	Quindi se nn riesci a chiamarmi prova l altro nume...	chat006	10
5389	13:10	24/dic/2014	Diletta	Dopo facciamo l albero e ti mando una foto	chat006	9
5473	21:52	24/dic/2014	Diletta	Ti piace l albero??	chat006	4
5721	20:40	26/dic/2014	Diletta	Come va l ape tesoro?	chat006	5
6334	19:08	07/gen/2015	Diletta	No nn sn stanca anche se nn ho dormito bene. Sopra...	chat006	25
7441	00:23	15/ott/2014	Matilda	ki nemmeno te l aspetto	chat011	5
7476	00:33	15/ott/2014	Matilda	ma spesso l odea ke noi abbiamo e diversa	chat011	9
8078	17:17	21/ott/2014	Isabell	In piu ci e stata l allusione	chat016	7
8137	11:24	25/ott/2014	Isabell	Non ci sn soldi in nero. niente per l a guardia di...	chat016	13
8280	11:58	08/nov/2014	Isabell	Per l italia?	chat016	3
8575	10:21	14/gen/2015	Isabell	A dopo, nn vedo l ora!	chat016	6
8590	10:54	14/gen/2015	Isabell	Nn ho l ombrello...	chat016	4

Abbildung 8: WhatsApp-Korpus – Ausschnitt aus Trefferliste bei Abfrage von fehlendem Apostroph bei bestimmtem Artikel (Katharina Jakob)

Das Korpus dokumentiert einen attraktiven Bereich der sprachwissenschaftlichen Lehre; es wird zweifellos schnell zum Gegenstand anderer studentischer Arbeiten werden, die es dann nicht nur verwenden,

sondern gleichzeitig weiter ausbauen und womöglich differenzierter strukturieren sollen.

Forschungskommunikation dient der Aufklärung

Digitale Forschungskommunikation dient also, genauer gesagt, der digitalen Aufklärung, dem eigentlichen Unterrichtsziel unseres Studiengangs. Frei nach der berühmten Definition des preußischen Aufklärers Immanuel Kant⁵ ist digitale Aufklärung der Ausgang des Geisteswissenschaftlers aus seiner selbstverschuldeten digitalen Unorganisiertheit. Unorganisiertheit ist das Unvermögen, sich seiner eigenen Forschungsdaten ohne Abstimmung mit den Daten anderer zu bedienen. Selbstverschuldet ist diese Unorganisiertheit, wenn die Ursache derselben nicht am Mangel der Forschungsdaten, sondern der Bereitschaft liegt, sich seiner eigenen Daten ohne Abstimmung mit den Daten der Kommilitonen und Fachkollegen zu bedienen.

Literatur

VerbaAlpina: <https://www.verba-alpina.gwi.uni-muenchen.de/>

Aristotle (1933): *Metaphysics*, Volume I. Translated by Hugh Tredennick, Loeb Classical Library 271. Harvard.

Dylan, Bob (1965): *Ballad of a thin man*, Highway 61 revisited. <https://bobdylan.com/songs/ballad-thin-man/> (04.01.2018).

Kant, Immanuel (1784): *Beantwortung der Frage: Was ist Aufklärung?*, *Berlinische Monatsschrift*, 481–494. https://www.rosalux.de/fileadmin/rls_uploads/pdfs/159_kant.pdf (04.01.2018).

Kant, Immanuel (1784): *Beantwortung der Frage: Was ist Aufklärung?*, *Berlinische Monatsschrift*, 481–494. https://www.rosalux.de/fileadmin/rls_uploads/pdfs/159_kant.pdf (04.01.2018).

5 Die Vorlage des folgenden Passus lautet im Original: «Aufklärung ist der Ausgang des Menschen aus seiner selbstverschuldeten Unmündigkeit. Unmündigkeit ist das Unvermögen, sich seines Verstandes ohne Leitung eines anderen zu bedienen. Selbstverschuldet ist diese Unmündigkeit, wenn die Ursache derselben nicht am Mangel des Verstandes, sondern der Entschließung und des Mutes liegt, sich seiner ohne Leitung eines andern zu bedienen.» (Kant 1784, 481)

- Krefeld, Thomas (2017b): Kooperation. In: VerbaAlpina-de 17/1, Methodologie. https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&letter=K&db=171-22 (04.01.2018).
- Krefeld, Thomas (2017c): Wissenschaftskommunikation im Web. In: VerbaAlpina-de 17/1, Methodologie. https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&db=171&letter=W-62 (04.01.2018).
- Krefeld, Thomas & Lücke, Stephan (2017): Nachhaltigkeit – aus der Sicht virtueller Forschungsumgebungen, Korpus im Text. <http://www.kit.gwi.uni-muenchen.de/?p=5773&cv=7> (04.01.2018).
- Lücke, Stephan (2017): Digitalisierung. In: VerbaAlpina-de 17/1, Methodologie. https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&db=171&letter=D-15 (04.01.2018).
- McLuhan, Marshall (1962): The Gutenberg Galaxy. London.
- Montaigne, Michel de (1595): Les Essais, <http://www.bribes.org/trismegiste/montable.htm> (04.01.2018).