



Introduction

PÄIVI PAHTA, MERI LARJAVAARA
& CHRISTOPHER SCHMIDT

Introduction

Neuphilologische Mitteilungen, the peer-reviewed scholarly journal founded by the Modern Language Society of Helsinki in 1899, publishes philological and linguistic research on German, English and Romance languages. Starting from the current issue of the 121st volume, *Neuphilologische Mitteilungen* becomes an Open Access digital publication, and ceases to be distributed on paper. To celebrate this significant historical step promoting open science, several articles in a thematic section of the current issue address the new modalities of research that are also relevant for the transformation of the journal: digitization and open access. The studies provide insight to the new opportunities for empirical research given by these new modalities while at the same time giving rise to more theoretical reflections on their impact on study of languages.

In the humanities and more specifically in the language sciences during the second decade of our century, the key words for the practical modalities of research have been digitization and open access. In many areas of research, these two phenomena are becoming *sine qua non* conditions for scholarly work; at the same time, it goes without saying that they push the boundaries of research and re-shape its contours.

The methods linked to and enabled by digitization open new paths and raise new scientific questions. For example, let us think of work based on different types of megacorpora, such as *Frantext* (over 250 million words), *Hansard Corpus* (1.6 billion words), or *iWeb* (14 billion words), or of studies on the linguistic features occurring in texts representing different genres (Biber & Egbert 2018, Laippala et al. 2019, Legallois et al. 2018). The policies of open science aim at providing open access to research data and methods, thus making them available to researchers on a much broader scale than before. Open access to research publications, on the other hand, allows researchers to take a quick look, without extensive bibliographic searches, at studies that at first glance do not seem central to the work in progress, but, on closer inspection, can contribute relevant information. They can do it wherever they are, and without licensed access to publishers' or institutions' digital resources. Furthermore, digitization and open access both contribute to opportunities for pluri- and transdisciplinarity.

In the long run, new research methods can lead to changes in the scientific paradigm. We want to raise the question of whether this has already happened in linguistics and philology following the emphasis on digitization and open access. Time will tell.

Introduction

Neuphilologische Mitteilungen, revue fondée en 1899 par la Société Néophilologique de Helsinki et recouvrant les domaines philologique et linguistique de l'allemand, de l'anglais et des langues romanes, sera à partir de ce numéro en libre accès et cessera d'être imprimée sur papier. Pour fêter ce grand pas en avant, le numéro contient plusieurs articles autour d'une thématique commune : les nouvelles modalités de recherche, la numérisation et le libre accès. Certaines études dans ce numéro profitent de corpus disponibles sur Internet, d'autres portent un regard plus théorique sur les nouvelles possibilités.

En sciences humaines et plus particulièrement en sciences du langage, la deuxième décennie de notre siècle a vu apparaître de nouvelles modalités pratiques de recherche comme justement la numérisation et le libre accès. Pour une partie des chercheurs, ces deux phénomènes sont en train de devenir les conditions *sine qua non* de toute recherche qui se respecte. En même temps, il est clair qu'ils tracent les limites de la recherche et lui dessinent de nouveaux contours.

Les méthodes liées à la numérisation et permises par celle-ci ouvrent de nouvelles voies et font surgir de nouveaux questionnements scientifiques : pensons par exemple aux travaux sur différents types de corpus gigantesques (*Frantext* avec 258 millions de mots, *Hansard corpus* avec 1,6 milliards de mots, *iWeb* avec 14 milliards de mots, entre autres) ou aux études sur les traits linguistiques de différents genres de textes (telles que Biber & Egbert 2018, Laippala *et al.* 2019, Legallois *et al.* 2018). Le libre accès de la recherche, pour sa part, permet aux chercheurs d'avoir un accès aisé et gratuit, où qu'ils soient, à de grandes quantités de données et de méthodes mais également de jeter un coup d'œil rapide, sans recherches bibliographiques extensives, à des études qui au premier abord ne semblent pas centrales pour le travail en cours, mais

qui peuvent y apporter des éléments tout à fait intéressants. La numérisation et le libre accès soulignent tous les deux la pluri- et la transdisciplinarité.

À la longue, les nouvelles modalités de recherche peuvent engendrer des changements de paradigme scientifique. La question de savoir si cela a déjà été le cas en sciences du langage et en philologie suite à l'accent mis sur la numérisation et le libre accès se pose légitimement. L'avenir nous le dira.

Vorwort

Die *Neuphilologischen Mitteilungen* ist eine Zeitschrift, welche im Jahr 1899 durch den Neuphilologischen Verein in Helsinki gegründet wurde und nebst philologischen auch die linguistischen Forschungsbereiche in den deutschen, englischen und romanischen Sprachen abdeckt. Ab dem Jahr 2020 geht die Zeitschrift zum Publikationssystem des *Open Access* über, weshalb ab diesem Jahr die Printversionen aufhören werden. Um diesen bedeutenden Schritt angemessen würdigen zu können, befasst sich diese erste *Open-Access*-Ausgabe in einem gesonderten Teil mit den Möglichkeiten und Tendenzen der Digitalisierung in der philologischen Forschung.

Im Bereich der Geisteswissenschaften und insbesondere in den Sprachwissenschaften haben Digitalisierung und *Open Access* immer mehr an Bedeutung gewonnen. Für manche Forscher sind diese zwei Phänomene zur Grundvoraussetzung wissenschaftlichen Arbeitens geworden; gleichzeitig werden dadurch die Rahmenbedingungen der Forschung neu abgesteckt und das Profil wissenschaftlichen Arbeitens verändert.

Die Methoden, welche sich aus der Digitalisierung ergeben und durch sie ermöglicht werden, können neue Forschungswege eröffnen sowie auch neue wissenschaftliche Fragestellungen ermöglichen. Dies betrifft unter anderem solche Bereiche wie das Arbeiten mit umfangreichen Korpora oder das Bearbeiten vielseitiger sprachlicher Phänomene in den verschiedensten Textsorten. Hier können z.B. solche Korpora wie Frantext (über 250 Mio. Wörter), das Hansard Corpus (1,6 Mrd. Wörter) oder iWeb (14 Mrd. Wörter) genannt werden, aber auch viele andere mehr. Das Prinzip transparenter Wissenschaft dient dazu einen offenen Zugang zu Forschungsdaten und -methoden zu garantieren, wodurch Forschungsergebnisse einer weitaus größeren Forschungsgemeinschaft zugänglich werden.

Open Access-Publikationen wiederum erlauben den Wissenschaftler*innen einen zeitnahen und unkomplizierten Zugang zu anderen Forschungsergebnissen; auch solchen, die auf den ersten Blick vielleicht für laufende Projekte nicht zentral zu sein scheinen, dessen ungeachtet dennoch einen interessanten Mehrwert bieten können. Digitalisierung und Open Access verleihen der wissenschaftlichen Inter- und Transdisziplinarität eine neue Dimension.

Digitalisierung von Daten und digitale Zugangssysteme zu Daten und ihrer Verarbeitung äußern sich jedoch nicht nur durch ihre offene Thematisierung, sondern begleiten Wissenschaft täglich im Hintergrund, sozusagen als unsichtbare Hand; nicht minder mit entscheidender qualitativer Zubringerfunktion zu wissenschaftlichen Ergebnissen. Um diesen Umstand zu würdigen, veröffentlicht die vorliegende Ausgabe neben Beiträgen, die den Digitalisierungsprozess als solchen explizit thematisieren als auch solche, die in Fortsetzung der über 100-jährigen Tradition der Zeitschrift digitale Zugangsmethoden allenfalls implizit erkennen lassen, jedoch nicht minder durch sie geprägt werden.

Auf lange Sicht können neue Forschungsmethoden auch zu einem Paradigmenwechsel in der Wissenschaft führen. Inwieweit die Digitalisierung von Forschungsinhalten und -methoden die Linguistik und Philologie prägen und neu verorten kann, wird die Zukunft zeigen.

PÄIVI PAHTA

MERI LARJAVAARA

CHRISTOPHER SCHMIDT

Corpora – Corpus

Frantext www.frantext.fr

Hansard Corpus <https://www.english-corpora.org/hansard/>

iWeb <https://www.english-corpora.org/iweb/>

References – Références

BIBER, Douglas – Jesse Egbert 2018.
Register Variation Online. Cambridge:
Cambridge University Press.

LAIPPALA, Veronika – Roosa Kyllönen –
Jesse Egbert – Douglas Biber – Sampo
Pyysalo 2019. Toward Multilingual
Identification of Online Registers.
*Proceedings of the 22nd Nordic
Conference on Computational
Linguistics*. NEALT Proceeding Series
42. 292–297. [https://www.aclweb.org/
anthology/W19-61.pdf](https://www.aclweb.org/anthology/W19-61.pdf)

LEGALLOIS, Dominique – Thierry
Charnois – Meri Larjavaara éds. 2018.
*The Grammar of Genres and Styles:
From Discrete to Non-Discrete Units*.
*Trends in Linguistics – Studies and
Monographs* 320. Berlin – Boston: De
Gruyter Mouton.