

/instituut voor de Nederlandse taal/

(English below)

Stageplaats t.b.v. De digitale Pallas

We zoeken een stagiair die de 300 talen die in het Russische *Vergelijkende woordenboek van alle talen en dialecten* (Сравнительный словарь всех языков и наречий, по азбучному порядку расположенный) (1790-1791) door Peter Simon Pallas worden genoemd, te identificeren op basis van *Ethnologue* (<https://www.ethnologue.com/>): wat is de moderne Engelse naam, tot welke taalfamilie wordt de taal tegenwoordig gerekend en waar op de wereld wordt de taal gesproken (die informatie willen we koppelen aan een wereldkaart zodat inzichtelijk wordt hoe de beschreven talen zijn verdeeld over de wereld). Kennis van vergelijkende taalwetenschap is daarbij een voordeel. Het schrijven van een onderzoekspaper op basis van de gegevens kan onderdeel zijn van de stage.

Contact en begeleiding: [Nicoline van der Sijs](#)

Opdracht: De 300 talen in het Russische *Vergelijkende woordenboek van alle talen en dialecten* door Peter Simon Pallas zijn in het Russisch vermeld. Van sommige grotere talen zijn de Engelse namen al bekend en toegevoegd, maar er is ook sprake van veel inheemse talen in o.a. Noord- en Zuid-Amerika, Afrika en het grote Russische rijk, die met verouderde namen worden aangeduid of met namen die slechts verwijzen naar de plaats waar ze toentertijd werden gesproken. We zoeken een stagiair die de Russische taalnamen kan lezen en die bereid en in staat is via die naam en informatie uit secundaire, veelal Duitstalige literatuur, de door Pallas genoemde talen zo goed mogelijk in kaart te brengen.

Als de verschillende talen zijn geïdentificeerd is de volgende stap om te achterhalen hoe Pallas aan zijn informatie over deze talen is gekomen. Van sommige talen weten we dat hij gebruik gemaakt heeft van een woordenboek, maar welk?

/instituut voor de Nederlandse taal/

Achtergrond: Op initiatief van de Russische tsarina Catharina de Grote stelde de Duitse onderzoeker Peter Simon Pallas een programma en vragenlijst samen om materiaal van de talen van Rusland en de hele wereld te verzamelen ten behoeve van taalvergelijkend onderzoek. De vragenlijst werd niet alleen verstuurd naar de administrateurs van de Russische provincies, maar ook naar Europa, China, Brazilië en Noord-Amerika. Op basis van het verzamelde materiaal publiceerde Pallas een vergelijkend woordenboek in twee delen (1787-1789). Het eerste deel bevat 285 concepten in 149 talen. Het tweede deel bevat telwoorden in 222 talen.

In 1790-1791 werd het woordenboek opnieuw uitgegeven in vergrote en gewijzigde vorm. De nieuwe vierdelige editie presenteerde bijna 300 talen en dialecten. Volgens het voorstel van Catherine werd de structuur van het woordenboek gewijzigd: de woorden in de verschillende talen van de wereld waren nu alfabetisch gerangschikt.

Het materiaal is nog steeds erg belangrijk voor de vergelijkende taalkunde en de historische taalkunde, en bevat ook data van bedreigde talen. Het materiaal kan allerlei nieuwe taalkundige en historische onderzoeksvragen beantwoorden en bijvoorbeeld worden gebruikt voor lexicostatistisch en colexificatie-onderzoek (zie <https://clics.clld.org/>). Tot nu toe heeft er nog maar relatief weinig onderzoek plaatsgevonden met het materiaal, omdat de gegevens niet zijn gedigitaliseerd en omdat ze in het Russisch en in het Cyrillische alfabet zijn geschreven.

In de afgelopen maanden zijn alle gegevens uit het woordenboek in een database ingevoerd, wat geleid heeft tot ruim 60.000 woorden uit 300 talen. Woorden, talen en begrippen kunnen via de database gemakkelijk gesorteerd worden.

Het project wordt gezamenlijk uitgevoerd door Tjeerd de Graaf, Wim Honselaar, Janine Jager, Sasha Lubotsky, Bruno Naarden en Vincent Wintermans.

Meer informatie over het woordenboek is te vinden op:

https://ru.wikipedia.org/wiki/Паллас,_Пётр_Симон

/instituut voor de Nederlandse taal/

Internship The digital Pallas:

We are looking for a trainee who can identify the 300 languages listed in the the Russian *Comparative dictionary of all languages and dialects* (Сравнительный словарь всех языков и наречий, по азбучному порядку расположенный) by Peter Simon Pallas (1790-1791) on the basis of *Ethnologue* (<https://www.ethnologue.com/>): what is the modern English name, to which language family does the language belong according to our present knowledge, and where in the world is the language spoken? (We want to link that information to a world map so that it becomes clear how the described languages are distributed around the world.) Knowledge of comparative linguistics is an advantage. Writing a research paper based on the data can be part of the internship.

Contact and guidance: [Nicoline van der Sijs](#)

Assignment: The 300 languages in the *Comparative dictionary of all languages and dialects* by Peter Simon Pallas are listed in Russian. The English names of some larger languages are already added, but there are many indigenous languages in, among others, North and South America, Africa and the great Russian Empire, which are indicated by obsolete names or with names that merely refer to the place where they were spoken at the time. We are looking for a trainee who can read the Russian language names and who is willing and able to find background information from secondary, mostly German-language literature, on the languages mentioned by Pallas.

Once the different languages have been identified, the next step is to find out how Pallas got his information about these languages. We know that for some languages he gathered his material from he a dictionary, but which dictionary?

/instituut voor de Nederlandse taal/

Background: On the initiative of the Russian Empress Catherine the Great, the German researcher Peter Simon Pallas compiled a program and a questionnaire for collecting material and comparative analysis of the languages of the peoples of Russia and the whole world. The questionnaire was sent not only to the administrators of the Russian provinces, but also to Europe, China, Brazil and North America. Having completed the study, Pallas published a comparative dictionary in two parts (1787-1789). The first volume included 285 concepts in 149 languages. The second volume contained the names of simple numerals in 222 languages.

In 1790-1791, the dictionary was reprinted in augmented and amended form. The new four-volume edition presented almost 300 languages and dialects. According to the proposal of Catherine, the structure of the dictionary was changed: the words in the various languages of the world were now alphabetically arranged.

The material is still very important for comparative linguistics and historical linguistics, and also presents data from endangered languages. The material can answer all kind of new linguistic and historical research questions, and for instance be used for lexicostatistic and colexification research (see <https://clics.cild.org/>), but it is as yet understudied because the data are not digitized and because they are written in Russian and in the Cyrillic alphabet.

In recent months, all data from the dictionary have been entered into a database, resulting in more than 60,000 words from 300 languages. Words, languages and concepts can be easily sorted within the database.

The project is jointly undertaken by Tjeerd de Graaf, Wim Honselaar, Janine Jager, Sasha Lubotsky, Bruno Naarden and Vincent Wintermans.

More information on the Dictionary is to be found at:

https://ru.wikipedia.org/wiki/Паллас,_Пётр_Симон
