

The Uppsala Russian Corpus

SND-ID: ext0071-1.

Skapare/primärforskare

Lennart Lönngren - Universitetet i Tromsø, Institutionen för språkvetenskap
Uppsala universitet, Institutionen för moderna språk

Forskningshuvudman

[Universitetet i Tromsø](#) - Institutionen för språkvetenskap

Beskrivning

Uppsalakorpusen består av runt 600 ryskspråkiga texter med sammanlagt en miljon ord (token), jämnt fördelade på informativ och skönlitterär prosa. De informativa texterna är skrivna 1985-1989 medan de skönlitterära texterna, vars vokabulär inte föråldras lika snabbt, hänför sig till en längre tidsperiod, från 1960-1988. Korpusen omfattar inte poesi eller drama.

Ett omfattande arbete har lagts på att säkerställa att korpusen blir så representativ och varierad som möjligt inom de givna ramarna. De informativa texterna är hämtade från 25 olika ämnesområden: ekonomi, utrikes nyheter/utrikespolitik, ideologi/inrikespolitik, partiangelägenheter, det sovjetiska samhället, sociala frågor, försvar, utbildning, lagstiftning, historia, kultur, lingvistik, medicin/hälsovård, psykologi, miljö/ekologi, jordbruk, ingenjörskonst, informationsteknologi, rymdforskning, energi, biologi, geologi/geografi, fysik, kemi och idrott. Vissa områden som ansågs vara viktigare representeras av en större textmängd.

Den skönlitterära hälften av korpusen omfattar verk av följande 40 författare: Abramov, Ajtmatov, Astaf'ev, Baklanov, Bek, Belov, Bitov, Bondarev, Dubov, Ganin, Gladyshev, Granin, Grekova, Goncharov, Iskander, Kaverin, Kazakov, Kochnev, Kozhevnikova, Nagibin, Lixanov, Lidin, Paustovskij, Pogodin, Pristavkin, Troepol'skij, Rasputin, Shcherbakova, Simonov, Solouxin, Shmelev, Tendrjakov, Tokareva, Tolstaja, Trifonov, Vasil'ev, Vorobl'ev, Zalygin och Zorin. Inte heller här är representationen jämn, utan de kändare författarna representeras av större textmängder.

För mer detaljerad information om korpusen se Lönngren, Lennart (ed.), 1993. Chastotnyj slovar' sovremennogo russkogo jazyka. (Modern rysk frekvensordbok. Med engelskspråkig sammanfattning.) Acta Universitatis Upsaliensis, Studia Slavica Upsaliensia 32. 188 sidor Uppsala. ISBN 91-554-3134-8.

Syfte:

Syftet är att tillhandahålla en korpus med ryskspråkiga prosatexter.

Data innefattar personuppgifter

Nej

Tidsperiod(er) som undersökts

1960 - 1988

1985 - 1989

Dataformat / datastruktur

[Text](#)

Resurstyp

Korpus

Ansvarig institution/enhet

Institutionen för språkvetenskap

Forskningsområde

[Språk och litteratur](#) (Standard för svensk indelning av forskningsämnen 2011)

[Språk och lingvistik](#) (CESSDA Topic Classification)

Nyckelord

[Ryska](#), [Korpus](#)

Publikationer

Lönngren, Lennart (ed.), 1993. Частотный словарь современного русского языка. [Chastotnyj slovar' sovremennogo russkogo jazyka.] (A Frequency Dictionary of Modern Russian. With a Summary in English.) Acta Universitatis Upsaliensis, Studia Slavica Upsaliensia 32. 188 pp. Uppsala
ISBN: 91-554-3134-8

Tillgänglighetsnivå

Åtkomst till data via extern aktör

Data är fritt tillgängliga

Hemsida

[Uppsala Russian Corpus](#)

CLARIN Virtual Collection Registry

[Lägg till i samling](#)

En virtuell samling är kopplad till ett specifikt forskningsändamål och innehåller länkar till dataresurser i olika digitala arkiv. Samlingen är lätt att skapa, få åtkomst till och citera.

Read more about virtual collections on the [CLARIN website](#).

Ladda ner metadata

[DataCite](#)

[MetaShare](#)

[MetaShare-CMDI](#)

[DDI 2.5](#)

[DDI 3.3](#)

[DCAT-AP-SE 2.0](#)

[JSON-LD](#)

[PDF](#)

[Citation \(CSL\)](#)