

TESTVEJLEDNING til ordforrådstestene

Derdiedas

1) De 5 test

Derdiedas 500

Dette er den første af i alt 5 ordforrådstests til tysk på de gymnasiale uddannelser i Danmark. Derdiedas 500 undersøger elevens kendskab til de hyppigste 500 ord på tysk. Med 500 ord kan ml. 65 og 80 % af teksten forstås, så det er et meget vigtigt grundlag.

Derdiedas 1500

Derdiedas 1500 testen undersøger, om eleven kender til de 1500 hyppigst brugte ord på tysk. Det vil være et passende ordforråd at have i starten af 1g, da det letter læsningen af gymnasiets tekster betragteligt. 1500 passer til det internationale niveau A2.

Derdiedas 2500

Derdiedas 2500 tester elevens kendskab til ordene i frekvensområdet 1500 – 2500, dvs. stadig de hyppige ord på tysk. Et ordforråd på omkring 2500 ord giver eleven mulighed for at læse og forstå tyske tekster med støtte i form af undervisning, glosser og ordbog; dvs. være i stand til at deltage i tyskundervisningen i 1g (og 2g). 2500 ord passer til det internationale niveau B1.

Derdiedas 3500 og derdiedas5000

De sidste ordforrådstest er beregnet til elever på højeste niveau. De tester ord i frekvensområdet 3000 – 5000. Hvis eleven klarer sig godt her, vil han/hun rimelig ubesværet kunne læse mange tyske tekster. 5000 passer til det internationale niveau B2, det danske A-niveau på STX/HHX.

Antagelserne om forbindelsen ml. elevernes læsekompetence og ordforrådsstørrelse baserer sig på undersøgelser fra andre sprog og fra egne projekter. Du kan læse om mine resultater, da jeg undersøgte forholdet ml. elevernes tyske ordforråd og læsekompetencer, på hjemmesiden derdiedas.dk

2) Testbeskrivelse

Testen er en multiple-choice test, som måler elevens receptive viden om testord på skrift. Testen er en videreudvikling af den internationalt anerkendte VLT (Vocabulary Levels Test), oprindeligt skabt af Paul Nation i 1980'erne, udgivet i 1990 og videreudviklet af Schmitt, Schmitt & Clapham i 2001. I den oprindelige test var der 6 ord, der skulle matches med 3 definitioner. I min udgave er der 8 tyske ord, der skal matches med 4 danske ord. Fordelen ved at kunne bruge L1 (dvs. elevernes modersmål) i stedet for definitioner er, at det kræver et yderligere ordforråd at forstå definitioner på L2 (tysk), hvilket skaber støj i fortolkningen af

resultaterne. Jeg har yderligere udvidet testformatet fra 6 til 8 muligheder for at mindske risikoen for gætteri ved udelukkelsesmetoder, hvilket jeg stødte en del på i første runde med pilotforsøg.

Testen kan med stor sikkerhed udtale sig om, hvor stort elevens receptive, skriftlige ordforråd er. Den kan ikke udtale sig om elevens produktive ordforråd samt om elevens lyttekompetence.

Det praktiske¹

Derdiedastestene foregår på papir med blyant og uden brug af hjælpemidler.

Der findes også online-udgaver på hjemmesiden derdiedas.dk. Disse tests er baseret på de originale tests og indeholder samme testord samt distraktorer. Forløbet med kvalitetssikring fandt sted med de analoge tests, men sammenligninger mellem elevers resultater analogt og digitalt viser ret stor overensstemmelse. Vær dog opmærksom på elevens adgang til google translate og ordbog ved onlinetests.

Læreren må sørge for, at eleven har forstået opgaven ved at læse indledningen op, men må ellers ikke hjælpe undervejs. En hel klasse kan sagtens tage testen samtidig, bare de ikke kan kigge hos hinanden.

Hver test tager ca. 10 – 15 minutter, men der er ingen tidsgrænse. Har eleven behov for at læse ordene langsomt eller gennemgå testen flere gange, er det helt o.k. I den ideelle læsesituation er dette jo også tilfældet.

Testen rettes ved at se, om de rigtige bogstaver står uden for de rigtige tyske ord. Har eleven valgt at skrive fx A ud for to tyske ord, gælder det også som en fejl.

3) Korpus

Frekvenstanken

Testen er frekvensbaseret. Den letteste test derdiedas500 tester således elevens kendskab til ord, der har den højeste frekvens på tysk, dvs. forekommer hyppigst i tyske medier og udgivelser. Det baserer sig på den antagelse, der er blevet bekræftet utallige gange i både tyske og andre forskningsresultater, at de ord, som man møder oftest, gennemsnitligt vil være de ord, som man tilegner sig hurtigst.

Den procentvise andel af ord, som den enkelte elev kender i hvert frekvensniveau, giver et samlet billede af elevens ordforrådsniveau. De enkelte ord er ikke interessante i sig selv, men det de repræsenterer, nemlig et frekvensniveau, er interessant. Kender en elev et givent ord fra et frekvensniveau, er der stor chance for, at den samme elev nok også kender andre ord fra samme frekvensniveau.

¹ I forsøgsperioden var jeg til stede under testen, og det var også mig, der rettede dem efterfølgende.

Kender eleven således 80 % af ordene i *derdiedas500*, kan den pågældende elev og hans/hendes lærer gå ud fra, at denne elev har et ret godt kendskab til de hyppigste 500 ord på tysk.

Frekvenslisten er baseret på lemma'er, hvilket er et korpuslingvistisk fagudtryk. Et lemma er et ord i sin grundform. Ordene *lesen, lese, liest, lesen, lest, las, last, lasen og gelesen* er således kun medregnet som ét lemma, nemlig verbet „lesen“, mens fx substantivet „Der Leser“ er medtaget som et andet lemma. Man kan selvfølgelig - med rette - sætte spørgsmålstegn ved, om eleverne kender alle verbalformer af netop de uregelmæssige verber, så det skal understreges, at netop disse må trænes og evt. testes i en anden sammenhæng. Hvis frekvenslisten havde indeholdt alle grammatiske former på tysk (fx også substantivers genitiv og pluralformer), ville det til dels være svært at sammenligne med erfaringer fra andre sprog og til dels ville antallet af ord, der skulle læres, kunne taget modet fra enhver.

Udvælgelse af testord og distraktorer

Jeg har skabt mit eget korpus på baggrund af Sketch Engine's 16 milliarder store samling af tyske ord. Jeg har opstillet ordene efter frekvens i et Excelark og dernæst kontrolleret hvert eneste ord. Jeg har dobbelttjekket i Sketch Engine's samlinger af eksempelsætninger for at sikre mig, at ordet registreres under den rigtige ordklasse, og jeg har tjekket "range" for de mere usædvanlige ord. Hvor frekvens henviser til, hvor hyppigt Sketch Engine har fundet ordet i tyske medier og udgivelser, henviser range til, hvor bred en udbredelse ordet har. Hvis ordets udbredelse har været for smal, dvs. godt nok bruges ofte, men hele tiden kun på én type hjemmeside, er ordet stadig en del af frekvensordlisten, men de spiller ikke en rolle i testene. Det er f.x. tilfældet med en række ord til alt det juridiske og tekniske omkring det at have en hjemmeside. Det gælder ord som "Benutzer", "Anmeldung", "zustimmen" og "unverbindlich". Jeg har også haft min tvivl omkring ord fra sportens og kriminalitetens verden, men da mange gymnasieklasser beskæftiger sig med netop disse to emner, har hyppige specialudtryk fra disse fået lov at blive i testen.

Ca. 10 % af ordene i korpus er identiske med danske ord, eksempelvis "Hund". Da testen ikke vil virke troværdig overfor eleverne, hvis de møder disse ord i testen, blev disse ord udelukket fra at være testord og distraktorer. Disse 10 % kan man til gengæld med rette italesætte som de **10 % gratis ord**, som danske elever får, når de lærer tysk – bare fordi de er danskere 😊.

Andre ord, som optræder på frekvenslisten, men ikke bruges som testord er: kollokationer, proprier, forkortelser samt låneord fra andre sprog.

Ca. 15 % af ordene i korpus er med et engelsk ord "close cognates". De ligner hinanden så meget, at man ofte vil kunne gætte betydningen uden at kende det tyske ord, fx *Handlung/handling* og *Sicherheit/sikkerhed*. Da det jo ikke er snyd at være dansker, medtages disse ord i testen. De udgør – som i korpus - gennemsnitlig 15 % af testordene, så testen afspejler den virkelige verden.

I de to højeste tests har jeg også medtaget 10 % sammensatte ord i testen, da de er en del af frekvensniveauet 4000 – 5000. Det kan spekuleres i, om inddragelsen af sammensatte ord på det punkt gør disse tests 10 % lettere. Det vil fremtidens testevalueringer vise.

Alle ordene er desuden tjekket for gengangere. Det er vigtigt, at en elev oplever en frisk start i hver testbox og i princippet også fra testniveau til testniveau. Hvis "vorwerfen" således optræder som testord ét enkelt sted, har jeg undladt at bruge "Vorwurf" som testord et andet sted.

4) Pilotforsøg og validitetstjek

Der har været flere trin i kvalitetssikringen. De beskrives her i kronologisk rækkefølge.

- En testrunde med sammensatte ord (komposita). Her undersøgte jeg 40 2g-elevs tilgang til sammensatte ord for at vurdere, om disse skulle inddrages i testen. Hvis et sammensat ord ikke blev opfattet som et samlet ord, men som 2 enkelte ord, ville jeg undlade at bruge det som testord. Det viser sig, at langt de fleste elever **ikke** oplever, at sammensatte ord er lettere at gennemskue end ikke-sammensatte ord. Det lader til, at de danske elever ofte har et for smalt basisordforråd til at gennemskue de to betydninger i det sammensatte ord, og af den grund tvinges til at opfatte sammensatte ord som ét samlet nyt ord, de skal lære. Af den grund medtages de sammensatte ord i testen.
- En testrunde med 2 forskellige testformater. En multiple-choice-test (MCT) med 3 testord og 3 distraktorer (a la VLT) gennemføres med 12 elever. En kontrolgruppe med 12 elever testes i en simpel oversættelsestest fra tysk til dansk med de samme testord. Da forskellen ml. de to grupper er synlig (til MCT-gruppens fordel), foretager jeg mundtlige interview med udvalgte MCT-elever for at undersøge, om deres bedre resultater skyldes u hensigtsmæssigt gætteri. Den gode nyhed er, at gættetiet oftest foretages på baggrund af kvalificeret viden om lignende ords betydning på tysk, men alligevel vælger jeg at udforme en ny MCT-test med 4 testord og 4 distraktorer for at mindske gættetiets betydning for elevens samlede resultat.
Herunder ser du resultatet af mine gætteri-undersøgelser efter brugen af det nye format.

Jeg har interviewet 10 testtagere for at se, om de i testen bliver krediteret for at kende ord, som de ikke kender i virkeligheden – med andre ord: hvor meget opfordrer testen til gætteri?

Svarer eleven rigtigt i testen og i interviewet, er der ingen problemer. Det samme gør sig gældende ved 2 x forkert svar. Hvis eleven svarer rigtigt i testen, men ikke kan oversætte ordet til dansk i et mundtligt interview, er der tale om overestimering, dvs. eleven har (måske) fået for mange point. Hvis eleven svarer forkert i testen, men kender ordet i virkeligheden, er der tale om underestimering og eleven har (måske) fået for få point.

Interview ↓ / Testsvar →	rigtigt	forkert
rigtigt	😊	underestimering
forkert	overestimering	😊

MCT-formatet vil altid opfordre til gætteri, og som nedenstående oversigt viser, foregår der også gætteri i derdiedas-testene. Især tallene for overestimering, dvs. der hvor eleven har "gættet" rigtigt uden at kunne oversætte ordet til dansk, er høje. Men der gemmer sig en positiv historie bag tallene. I lighed med Vilkaité, Hyllstad og Schmitts forskningsresultater fra 2015 er tallene for overestimering ganske vist høje, men når man får et indblik i elevernes strategier ved at bede dem om at tænke højt, mens de gennemfører testen, ser man, at deres strategier indeholder megen god viden om fremmedsproget. Mine forsøgspersoner kunne ofte genkende ordets betydning pga. ligheder med dansk, engelsk eller andre tyske ord, eller de kunne genkende dele af ordet og dermed gætte betydningen. Men den hyppigst anvendte strategi var nok udelukkelsesmetoden, hvor de kendte de øvrige tyske ord, distraktorerne, men da disse ord stammer fra samme frekvensområde, er det jo lige så godt som at kende testordene. Af den grund er jeg ikke bekymret over elevernes "gætteri".

Resultater fra interviews:

	derdiedas500	derdiedas1500	derdiedas2500	derdiedas3500
underestimering (gennemsnit)	6 %	4 %	6 %	8 %
overestimering (gennemsnit)	11 %	17 %	22 %	29 %

- c) En analyse af distraktorer: hvor godt virker testenes distraktorer? En distraktor, der aldrig bliver valgt, er ikke særligt velvalgt. De skal helst kunne skabe tvivl hos de elever, der ikke kender testordene (for ellers fungerer de som en del af elevens udelukkelsesmetode), men samtidig må de heller ikke skabe usikkerhed hos de elever, der kender testordene. Distraktorerne skal altså have en tydelig anden betydning, men må heller ikke adskille sig så meget, at det er tydeligt, at de bare er med som distraktorer. Alle distraktorer stammer fra samme frekvensområde som testordene, de tilhører samme ordklasse, og det tilstræbes, at de har en sammenlignelig længde. Derudover må de ikke semantisk kunne skabe tvivl hos eleven ved at have en betydning, der ligger for tæt op ad testordet.
- d) Pilotforsøg med alle fem nye MCT-tests i efteråret 2019. Indtil videre har jeg 3 deltagende gymnasier, 10 deltagende tyskhold på hhv. 1g, 2g og 3g-niveau. 1g- og 2g-holdene tester derdiedas500 - 2500, mens 3g-holdet tester derdiedas3500 og derdiedas5000; udvalgte elever i 1g har også taget de to svære tests for at se, om man allerede ud fra derdiedas500 - 2500 kan forudsige deres samlede ordforrådsniveau, og det ser lovende ud. Ligesom i den tidligere runde vil

enkelte elever deltog i mundtlige interviews efterfølgende for at tjekke, om de nu også kender de ord, som de har fået rigtige i MCT-testen.

- e) Læsetest. Afslutningsvis har 126 elever fra 2 gymnasier deltaget i et forsøg, hvor de både tog en ordforrådstest og en læsetest, så jeg kunne se, om resultaterne fra ordforrådstesten forudsiger elevens læsekompetence. Sammen med Norbert Schmitt gennemførte vi en statistisk t-test og fandt en ret høj korrelation på 0.733.

5) Fortolkning af resultater

Fortolkningen af lignende MCT-tests, både på engelsk og tysk, er typisk, at man opfatter en given ordforrådstest som bestået, hvis eleven har min. 80 % rigtige.

Da der ikke er ret mange danske elever, der består derdiedas 500, da vi ikke arbejder frekvensbaseret med tyskundervisningen i grundskole og på gymnasie, vælger jeg at lade resultater fra alle tests tælle sammen for at give en samlet vurdering af en elevs ordforrådstørrelse.

elev nr.	derdiedas 500 - antal rigtige	derdiedas 1500 - antal rigtige	derdiedas 2500 - antal rigtige	sammentælling af antal rigtige	Ordforråds-niveau ud fra top 2500	Formodet samlet ordforråd*
elev nr. 7	22 = 69 %	19 = 59 %	12 = 38 %	345 + 590 + 380	1315	ca. 2000

*de foreløbige resultater viser, at man med ret stor sikkerhed kan udtale sig om det samlede ordforrådsniveau ud fra alene de tre første tests, men det kræver flere empiriske undersøgelser.

6) Samarbejde og støtte

Testene er blevet til i et tæt samarbejde med Professor Norbert Schmitt, der er anerkendt i hele verden for sin viden om ordforrådstestning. Han har bidraget med nyttig viden og nødvendige spørgsmål fra idéfasen, gennem den vigtige kvalitetssikring og til slutproduktet. Han er imidlertid professor i engelsk ordforråd, så ansvaret for det tyske indhold i disse test hviler alene hos mig.

UDLØDNINGSMIDLER TIL UNDERVISNING 2018 har givet projektet økonomisk støtte. Projektet havde ikke været mulig uden denne støtte.

NCFE (Nationalt Center for Fremmedsprog) bidrog med støtte til et forudgående projekt om sammensatte ord på tysk. Det var også meget værdifuldt.

P.s. om kollokationer: Kollokationer optræder i frekvenslisten, men ikke som del af den overordnede ordforrådstest. En senere test af kun kollokationer kunne være en mulighed. Mht. placeringen af substantiver, der ofte optræder i den samme kollokation, har jeg fulgt følgende fremgangsmåde: substantivet "Verlauf" er oplyst med hyppigheden 733.704, men når ordets brug undersøges, stammer 387.512 af ordets (fremkomst) fra kollokationen "im Verlauf". Derfor placerer jeg substantivet "Verlauf" med hyppigheden 346.192 (733.704 – 387.512).