NÅR ORD IKKE BARE ER ORD

Leksikalsk densitet og inventar i skriftlig engelsk

Projekt Forskerspirer 2017

"The number of people without jobs
is getting smaller because the amount of
work places is getting bigger."

"The number of unemployed people
has been reduced because of
the growing number of work places."

"The decrease in the employment rate
is due to the job creation increase."

"The reduction of unemployment
through job cumulation."

MILENA SIMON CHRISTENSEN
ROSBORG GYMNASIUM OG HF, 3.U
30/10/2017
FAGOMRÅDE: HUMANIORA

Milena Simon Christensen
Rosborg Gymnasium og HF, 3.U
30-10-2017

Indhold

1. Indledning 2
2. Problemformulering 2
3. Afgrænsning 3
4. Teori 3
4.1 Definition af akademisk sprog og ordforråd 3
5. Metode 5
6. Relaterede forskningsresultater 5
7. Udførelse 6
7.1 Pilotforsøg: 6
7.2 Resultat af analysen 7
8. Budget 8
9. Videre forskning 8
10. Foreløbig konklusion 9
11. Tak 9
12. Litteraturliste 10
13. Bilag 12
13.1 Tabel 1: Samlet oversigt for L1- og L2-datasæt 12
13.2 Tabel 2: Hyppighedstabel for AWL-ord (L1-datasæt) 13
13.3 Tabel 3: Hyppighedstabel for AWL-ord (L2-datasæt) 17
13.4 Hyppighedstabeller for AWL-ord i de enkelte tekster (L1) 19
13.5 Hyppighedstabeller for AWL-ord i de enkelte tekster (L2) 25

Milena Simon Christensen
Rosborg Gymnasium og HF, 3.U
30-10-2017

1. Indledning

I løbet af mine snart 2,5 år på gymnasiet har jeg, selvom jeg har engelsk på A-niveau, kun skrevet to former for skriftlige afleveringer. Begge går under betegnelsen "analytical essay", og de adskiller sig hovedsageligt fra hinanden ved, at den ene tager udgangspunkt i en fiktionstekst, den anden i en sagprosatekst. Opgaveformuleringerne til disse opgaver lægger mere vægt på korrekt grammatik, indhold (analyse og fortolkning) og form (essaystruktur) end på sproglige elementer som leksis og stilleje.

Ser man på den nye læreplan[1] for STX A engelsk, gældende for 1.g-elever fra aug. 2017, opstiller den bl.a. som fagligt mål, at eleven skal kunne:

> *"-skrive længere, **nuancerede** og velstrukturerede tekster på engelsk om komplekse emner med høj grad af grammatisk korrekthed, beherskelse af skriftsproglige normer samt formidlingsbevidsthed"*

I forhold til mit forskningsprojekt må man hæfte sig ved ordet "nuancerede", idet dette peger i retning af tilegnelse af et stort ordforråd. I "Engelsk A, STX, Vejledning" uddybes "nuanceret" til at betyde, "at eleverne har et omfattende ordforråd og anvender det med præcision"[2] (om "talefærdighed"). I afsnittet om "skrivefærdighed"[3] specificeres det, at der bl.a. skal være fokus på "nuancerigdom", og at det således kræves, "at eleverne har et nuanceret ordforråd". I forlængelse heraf fremgår det, at engelsk STX A niveaumæssigt svarer til C1 i Den Fælles Europæiske Referenceramme for sprog[4], hvilket er næsthøjeste niveau lige under 'native speaker'. Kravene til ordforråd kan altså siges at være betydelige.

Samtidig kan mere end 25% af kandidatuddannelserne på de danske universiteter tages på engelsk[5]. Man kan følgelig undre sig over, om danske gymnasieelevers skriftlige engelsk overhovedet udvikles nok igennem den bestående skrivepraksis til, at de vil være i stand til f.eks. at tage en videregående uddannelse på engelsk, og formidle og forstå akademiske ideer på et akademisk engelsk – og om eleverne er ligeså gode til dette, som de ville være på deres modersmål. Derfor vil dette projekt sammenligne danske 3.g elevers sprog med engelsksprogede elever, som går på sidste år i high school, idet begge grupper netop står overfor at påbegynde en videregående uddannelse og dermed kan forventes at have nået et højt sprogligt niveau.

2. Problemformulering

På baggrund af min egen erfaring fra engelskundervisningen, den nye læreplan for STX engelsk A-niveau og Dansk Sprognævns statistik har jeg valgt en problemformulering, der lyder: Hvordan kan en sammenligning af danske gymnasieelevers og modersmålstalendes skriftlige engelsk bidrage til en mere bevidst didaktik i forhold til densitet og stilleje i det leksiskalskeinventar?

---

[1] Læreplan, august 2017: https://uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017
[2] Engelsk A, STX, Vejledning, s. 6, afsnittet om talefærdighed. https://uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017
[3] Engelsk A, STX, Vejledning, s. 7, afsnittet om talefærdighed. https://uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017
[4] Ibid.
[5] Dansk Sprognævns rapport; "Dansk sprogs status 2012", s. 10.

3. Afgrænsning

Projektet lægger op til ændringer i engelskundervisningens didaktik, men jeg fokuserer på korpuslingvistisk analyse af tekster, og kommer derfor ikke ind på didaktisk teori.

Jeg analyserer analytiske essays skrevet til studentereksamen 2017 på engelsk A-niveau af danske gymnasieelever (STX), der har dansk som modersmål og engelsk som første fremmedsprog. Disse tekster sammenligner jeg med analytiske essays skrevet i 2017 af engelske high school-elever på sidste år, der har engelsk som modersmål. Jeg udfører kun produktanalyse, ikke analyse af den undervisning eleverne modtager, eller hvordan deres skriveprocesser forløber.

Herudover undersøger jeg ikke elevernes evne til at læse og forstå akademiske tekster, men kun hvorvidt deres engelsk skriftsprog er tilstrækkeligt akademisk til, at de vil være i stand til at formidle akademiske ideer. Derfor fokuserer jeg på, om der i teksterne er akademisk sprog og ordforråd, og altså ikke på f.eks. syntaks eller om teksterne er grammatisk korrekte. Dog undersøger jeg om bestemte akademiske ord anvendes korrekt i deres kontekst.

4. Teori

Projektet bygger på teori om, hvilke træk der er kendetegnende for akademiske tekster og hvordan disse kan findes i en tekst. Først og fremmest defineres akademisk sprog og ordforråd med udgangspunkt i forskellige teoretikere, og herefter beskrives konkrete mål- og analysérbare sproglige parametre.

4.1 Definition af akademisk sprog og ordforråd

For at kunne karakterisere akademisk ordforråd, må man skelne mellem det før-faglige, det faglige og det akademiske ordforråd. For at et ord kan indgå i det akademiske ordforråd, må det ikke være et af de 2000 mest frekvente ord i sproget (Henriksen 2013). Kendetegnende for akademisk ordforråd er også – i modsætning til det faglige ordforråd – at det går på tværs af de forskellige akademiske hovedområder (Nation 2013, Henriksen 2013).

Eftersom der ikke findes nogen fælles og simpel definition af akademisk sprog (Snow og Uccelli 2008), bruger jeg Nagy og Townsends (2012):

> *"Academic language is the specialized language, both oral and written, of academic settings that facilitates communication and thinking about disciplinary content".*

Akademisk sprog er altså ifølge Nagy og Townsend både mundtligt og skriftligt, men de skriver også, at man skriftligt vil se en større koncentration af de træk, som er typiske for akademisk sprog, end man vil i mundtligt sprog. Ifølge definitionen findes akademisk sprog især når man formidler fagligt indhold. Herunder nævner Nagy og Townsend 6 træk, som er typiske for akademisk sprog:

1. Latinsk og græsk ordforråd
2. Morfologisk komplekse ord
3. Substantiver, adjektiver og præpositioner
4. Grammatiske metaforer
5. Informationstæthed
6. Abstrakt sprog

I det følgende beskrives de ovenfor nævnte træk samt hvordan disse kan findes i en tekst, nærmere[6]:

Latinsk og græsk ordforråd:
Den større forekomst af latinsk og græsk ordforråd skyldes, at latin historisk har været det akademiske sprog i England, og at der derfor i mange akademiske sammenhænge ikke findes germanske ord med tilsvarende betydning (Nagy og Townsend). Hvorvidt en engelsk tekst har et stort akademisk ordforråd, kan måles med Coxheads "Academic Word List"[7], der – selvom den ikke udelukkende indeholder ord af latinsk og græsk oprindelse – indeholder de mest hyppigt brugte ord på tværs af fagområder (Coxhead 2000).

Morfologisk komplekse ord
Akademisk sprog indeholder længere ord og altså flere flerstavelsesord, hvilket er et resultat af præfikser og suffikser. Ofte bruges suffikser til at lave deriverede former[8]. Et sammenligneligt udtryk for dette ville man få ved at måle den gennemsnitlige ordlængde i en tekst eller et korpus.

Substantiver, adjektiver og præpositioner
Ser man på ordklasser, er en stor forekomst af substantiver mere kendetegnende for akademisk sprog end for ikke-akademisk sprog. Der er dog også i akademiske tekster flere adjektiver og præpositioner.

Grammatiske metaforer
Kort beskrevet bruges en grammatisk metafor til at formidle abstrakte teorier, processer eller idéer. En grammatisk metafor findes f.eks. i tilfælde, hvor noget abstrakt udfører en handling. Brugen af nominalisering er kendetegnende for grammatiske metaforer, idet omdannelsen af et verbum eller adjektiv til et substantiv netop muliggør, at abstrakte begreber kan være subjekt i en sætning[9].

Informationstæthed
Med informationstæthed menes, at akademiske tekster er informationsmæssigt mere komprimerede og dermed svære at forstå. informationstæthed er dog ikke målbar i sig selv, men som udtryk for tekstens sværhedsgrad kan man måle leksikalsk densitet. Leksikalsk densitet er defineret som mængden af indholdsord i forhold til den totale mængde ord (Mortensen 2006). Indholdsord defineres som substantiver, adjektiver, verber og adverbier. I forhold til leksikalsk densitet skal bemærkes, at *"mundtlige tekster ligger under 40% og skriftlige tekster over 40%"* (Ure 1971, citeret fra Mortensen 2006).

Abstrakt sprog
At akademiske tekster er mere abstrakte end andre tekster kan hovedsageligt ses på brugen af abstrakte substantiver, der f.eks. henviser til et koncept eller en proces fremfor en handling eller noget fysisk.

[6] Al teori i de nedenstående afsnit stammer fra Nagy og Townsend, 2012, medmindre andet er angivet.
[7] For Coxheads "Academic Word List" se https://www.victoria.ac.nz/lals/resources/academicwordlist/awl-headwords.
[8] Eksempel: *act > active, active > activate, activate > activation* (Nagy og Townsend 2012).
[9] Eksempel: *The package was heavy > The heaviness of the package* (Nagy og Townsen 2012).

5. Metode

Med udgangspunkt i teori om akademisk sprog og ordforråd har jeg valgt at arbejde med en sammenlignende empirisk analyse for at være så tæt på praksis som muligt.

Jeg har valgt at lave en synkron undersøgelse af teksterne, men det kunne have været en mulighed at lave en diakron undersøgelse, hvor man sammenligner niveauet i studentereksamensstile på to forskellige tidspunkter, eller man kunne følge en klasse fra 1.g til 3.g (en longitudinal undersøgelse), for at se hvilken effekt undervisning har. En synkron undersøgelse giver kun et aktuelt billede af en gruppe elevers præstation her og nu, men til gengæld muliggør den, at man kan undersøge en større population – og dermed nå frem til repræsentative resultater.

I udvælgelsen af empiri bygger jeg på de principper, der gælder i korpuslingvistikken. Det vil altså sige, at repræsentativitet og korpusstørrelse spiller en rolle (Mcenery et al. 2006). Gymnasieelevers essays vil være, hvad man kalder et specialiseret korpus (i modsætning til et referencekorpus), hvilket vil sige, at mens referencekorpuset bliver vurderet på "sampling", skal mit korpus vurderes på mætningsgrad: Hvor mange tekster skal der samles, for at det er sandsynligt, at ekstra tekster ikke vil tilføje ny information? Mortensen (2006) arbejder f.eks. med 111 essays i en lignende undersøgelse, og hvis man sammenholder dette med det antal elever, der skrev STXA-opgaven til eksamen i 2016[10], 19.216, skal mine korpora indeholde mellem 100 og 150 tekster hver for at være repræsentative. I valget af tekster til begge korpora er det også vigtigt, at teksterne er af nogenlunde samme længde, for at man kan sammenligne deres leksikalske densitet. Dette skyldes, at længere tekster har tendens til at have lavere leksikalsk densitet end kortere tekster (Johanson, 2008).

I forhold til dette projekt overvejede jeg også, om jeg skal annotere korpus eller ej. Ud fra hvad Mortensen (s. 15) skriver, har jeg valgt ikke at annotere manuelt, men derimod at anvende forskellige programmer til identificering af akademiske ord, dominerende ordklasser og leksikalsk densitet, hvilket jeg redegør nærmere for under udførelse.

Hvad spørgsmålet om kvantitativ versus kvalitativ undersøgelse angår, har jeg valgt en kombination. Den kvantitative analyse giver repræsentative tal, som er sammenlignelige, mens den kvalitative analyse gør det muligt at gå i dybden med enkelte eksempler og få svar på spørgsmålet om, hvorvidt eleverne også formår at anvende de forekommende ord korrekt i en kontekst.

6. Relaterede forskningsresultater

Hans Klarskov Mortensen har undersøgt og sammenlignet den leksikalske densitet i danske 2.g-elevers tekster med britiske A-level studerendes tekster. Denne undersøgelse viste, at den leksikalske densitet i A-leveltekstерne gennemsnitligt var ca. 8 procentpoint højere end i de danske elevers tekster (Mortensen 2006). Det er dog endnu ikke undersøgt, hvilke akademisk sproglige forskelle der er mellem danske elever, der har engelsk som første fremmedsprog, og elever med engelsk som modersmål, hvilket er fokus for dette projekt.

[10] I mangel af data fra studentereksamen 2017 benyttes tal fra studentereksamen 2016 fra rapporten: "Evaluering af skriftlig eksamen 2016 Engelsk STX/HF", s.7.

7. Udførelse

For at gøre det klart, hvordan jeg vil gå frem i min undersøgelse, beskrives i det følgende det pilotforsøg, som jeg har udført. Datamængden i dette forsøg er begrænset, og større korpora ville gøre undersøgelsens resultater mere valide. I et ideelt projekt ville jeg altså have større og mere repræsentative korpora. I pilotforsøget udgøres det empiriske grundlag af i alt 12 analytiske essays, og det antal tekster og den mængde tekst, der analyseres, er altså så lille, at der snarere er tale om datasæt end om deciderede korpora. Den praktiske udførelse ville dog være den samme i et forsøg med større datamængder som i det nedenfor beskrevne pilotforsøg.

I pilotforsøgets analyse har jeg valgt at fokusere på akademisk ordforråd målt med Coxheads AWL, de dominerende ordklasser og leksikalsk densitet og kommer derfor ikke ind på alle de ovenfor beskrevne træk[11], hvilket jeg ville gøre i udførelsen af et større forsøg.

7.1 Pilotforsøg:

I mit pilotforsøg var empirien to datasæt, hvoraf det første bestod af 6 analytiske essays skrevet på engelsk A-niveau til STX-afgangseksamen 2017 af danske gymnasieelever, mens det andet bestod af 6 analytiske essays skrevet af canadiske[12] elever. Det første datasæt kaldes L2[13] og det andet L1[14].

Datasæt L2 bestod af

1) Tekst A: karakter: 02
2) Tekst B: karakter: 12
3) Tekst C: karakter: 02
4) Tekst D: karakter: 12
5) Tekst E: karakter: 7
6) Tekst F: karakter: 7

Datasæt L1 bestod af

1) Tekst 1: karakter: A
2) Tekst 2: karakter: C
3) Tekst 3: karakter: D
4) Tekst 4: karakter: B
5) Tekst 5: karakter: C
6) Tekst 6: karakter: A

Som det ses er der en mindre forskel i forhold til karakterniveauet, men ikke en forskel, der er så stor, at den skønnes at være problematisk for en sammenligning.

Som første skridt i den kvantitative analyse af empirien fokuserede jeg på tre forskellige sproglige deskriptorer. Den første var optælling og udregning af procentvis mængde af ord fra Coxheads "Academic Word List"[15] i forhold til det totale antal ord i hver tekst. Den anden var udregning[16] af den procentuelle mængde af henholdsvis substantiver, adjektiver og præpositioner. Til sidst udregnedes den leksikalske densitet for den enkelte tekst og for de to datasæt som helhed[17].

[11] Se teoriafsnit.
[12] I mit pilotforsøg benyttede jeg canadiske essays skrevet af førsteårsstuderende, fordi det var det materiale, der var tilgængeligt.
[13] Language two – første fremmedsprog.
[14] Language one – modersmål.
[15] I praksis benyttes siden http://www.nottingham.ac.uk/alzsh3/acvocab/awlhighlighter.htm, som der henvises til fra den officielle AWL-hjemmeside, https://www.victoria.ac.nz/lals/resources/academicwordlist.
[16] Til udregning af procentuel mængde af de forskellige ordklasser samt leksikalsk densitet benyttede jeg programmet "Analyze My Writing": http://www.analyzemywriting.com/.
[17] Se Tabel 1 i bilag.

Andet skridt var en kvalitativ analyse og fortolkning af de kvantitative data. På baggrund af Nagy og Townsends påstand om, at det at kunne et ord betyder, at man er i stand til at anvende det i de rigtige sammenhænge, så jeg på, hvorvidt de forekommende AWL-ord blev brugt korrekt grammatisk og betydningsmæssigt.

Herefter foretog jeg en sammenligning af både kvantitative og kvalitative resultater fra L1-datasættet og L2-datasættet for at se, hvilke forskelle og ligheder der sprogligt var i teksterne, og dermed også hvilke områder man i særlig grad bør fokusere på i undervisningen for at forbedre danske gymnasieelevers engelske skriftsprog.

7.2 Resultat af analysen

Akademisk ordforråd målt med AWL:[18]

Denne del af analysen viser ikke en stor forskel på de to grupper. Gennemsnitligt er procentdelen af akademiske ord i L1-teksterne kun 1 procentpoint højere end i L2-teksterne. I L2-datasættet anvendes der 70 forskellige ord, mens eleverne i L1-datasættet anvender 127 forskellige ord. Forskellen på de to grupper er altså i dette tilfælde, at L1-eleverne lader til at have et større, mere varieret ordforråd. Herudover fremgår det også, at procentdelen af akademiske ord i L2-teksterne er uafhængig af den karakter opgaverne har fået.[19]

Substantiver, adjektiver og præpositioner:[20]

Hvad angår procentdelen af substantiver og adjektiver, er der større forskel imellem grupperne, idet L1-gruppen gennemsnitligt bruger henholdsvis 2,5 og 2 procentpoint flere af disse end L2-gruppen. På procentdel af præpositioner er der ingen forskel, idet begge sæt ligger på gennemsnitligt 13,5%.

Leksikalsk densitet:[21]

Ved leksikalsk densitet er der en væsentlig gennemsnitlig forskel på ca. 6 procentpoint. Denne forskel viser det samme som Mortensens undersøgelse[22]. I mit pilotforsøg ligger den L1-tekst, der har den højeste leksikalske densitet, på 57%, mens den L2-tekst, der har den højeste leksikalske densitet, ligger på 46,5%. Der er altså en forskel mellem disse to på over 10 procentpoint.

AWL-ord i kontekst:

De kvantitative data for AWL-ord og ordklasser viser ikke stor forskel på de to datasæt. I forhold til hvordan ordene anvendes, er der imidlertid stor forskel. I L2-gruppens essays gentages de samme ord oftere[23], hvilket gør, at sproget virker mindre nuanceret. Herudover er der i L2-sættet en stor forekomst af ord, der

[18] Se Tabel 1 i bilag.
[19] F.eks. er der både i tekst B og C 3% akademiske ord, men B har fået 12 og C har fået 02.
[20] Se Tabel 1 i bilag.
[21] Ibid.
[22] Se ”Relateret forskning”.
[23] I tekst B gentages ordet ”indicate” f.eks. 4 gange i et afsnit på 7 linjer.

benyttes tematisk[24] og dermed ikke i deres akademiske betydning. I L1-teksterne anvendes meget få akademiske ord i forkert kontekst, mens dette hyppigt ses i L2-teksterne[25].

I begge datasæt er nogle af de hyppigst benyttede ord[26], ord der i den kontekst de benyttes i bedre kan defineres som fagbegreber end som akademisk sprog[27]. Dette er et eksempel på, at ordene på Coxheads AWL, selvom de forekommer på tværs af videnskabelige hoved- og fagområder, kan have forskellige fagspecifikke betydninger (Coxhead 2016).

8. Budget

Selvom jeg i pilotforsøget har kunnet udføre analysen med gratis software, vil det i et større projekt være nødvendigt med indkøb af licenser til korpusværktøjer. Derfor har jeg i mit budget medregnet pris for korpusbehandlingsprogrammerne PhraseContext og Sketch Engine samt konsulenthjælp til disse.

Budget for projektet	
Rekvirering af eksamensopgaver fra danske elever: Besøg ved ca. 8 klasser à ca. 20 elever med rødvin til lærer og flødeboller til elever (estimeret)	2400 kr.
Rekvirering af essays fra britiske elever: Besøg ved ca. 8 klasser à ca. 20 elever med rødvin til lærer og flødeboller til elever (estimeret)	2400 kr.
Besøg på skoler i England	5000 kr.
Indkøb af licens til korpusværktøj: PhraseContext Sketch Engine	460 kr.
Konsulenthjælp til korpusværktøj	3500 kr.
Indkøb af ekstra sekundær litteratur (estimeret)	2500 kr.
Ekstern harddisk til lagring af tabeller og skemaer	800 kr.
Studentermedhjælper til korpusanalyse 145 kr. i 20 timer	2900 kr.
I alt	**19.960 kr.**

9. Videre forskning

Videre forskning bør undersøge, hvordan man kan udarbejde undervisningsmateriale ud fra de mangler i elevernes sprog, der klarlægges af dette projekt. Det ville her være relevant at undersøge vigtigheden af at give *"instruction without prescription",* noget, der er svært at gøre, hvis skriftligheden er baseret på genrer (Snow og Uccelli 2009), eller om man bør fokusere på brugen af akademiske ordlister i undervisningen

[24] I forhold til den tekst, der skrives om: f.eks. "sex", "adult" og "stability".
[25] I tekst A er der f.eks. en sætning, som lyder: "This comparison could also be interpreting as…". I dette tilfælde bruges verbet "interpret" i en forkert form, præsens participium (interpreting), hvor det skulle have været et perfektum participium (interpreted).
[26] Se Tabel 2 og Tabel 3 i bilag
[27] F.eks. "symbol", "theme" og "tense".

(Nation 2013). I forbindelse med dette ville det være interessant at lave en longitudinal undersøgelse, som kan vise om elevernes sprog forbedres gennem undervisning i akademisk sprog.

10. Foreløbig konklusion

Af pilotforsøget fremgår det, at L1-gruppen som forventet har et mere akademisk sprog end L2-gruppen, men de datasæt jeg har i mit pilotforsøg er for små til at vise, hvor stor forskellen er. Mønstre i datasættene viser dog allerede, at bl.a. leksikalsk densitet er noget, der adskiller danske elever væsentligt fra modersmålstalende. En måde at opnå høj leksikalsk densitet på er bl.a. ved nominalisering og deriverede former, hvilket er noget, man ville kunne fokusere på i undervisningen f.eks. ved at bevidstgøre om orddannelse i forbindelse med tekstlæsning eller direkte i undervisningsøvelser. Herudover er det ifølge undersøgelsen i høj grad nødvendigt at fokusere på ordforrådsindlæring samt korrekt kontekstuel og kollokational anvendelse af de indlærte ord. Undersøgelsen danner hermed grundlag for en bevidst fokuseret sproglig didaktik i engelskfaget.

11. Tak

Tak til min forskerkontakt, Birgit Henriksen, professor på Institut for Engelsk, Germansk og Romansk på Københavns Universitet, især for hjælp til litteratur og for henvisning til Anne Sofie Jakobsen, Ph.d. stipendiat ved Institut for Engelsk, Germansk og Romansk på Københavns Universitet, som har hjulpet med teori om akademisk sprog og ordforråd. Tak til Finn Nielsen, koordinator for Projekt Forskerspirer på Rosborg Gymnasium og HF, for vejledning og kage.

12. Litteraturliste

Bøger:

- Henriksen, Birgit: *Sprogsyn og læringssyn*. S. 21-34 i: Andersen, Hanne Leth et al: *Fremmedsprog i gymnasiet – teori, praksis og udsyn.* Frederiksberg: Samfundslitteratur, 2014.
- Hüttner, Julia Isabel: *Academic Writing in a Foreign Language*. Frankfurt am Main: Peter Lang, 2007.
- Mcenery et al.: *Corpus-Based Language Studies – an advanced resource book.* New York: Routledge Applied Linguistics, 2006.
- Mortensen, Hans J. Klarskov: *Når elever skriver engelsk.* Århus: Systime, 2006.

Videnskabelige artikler:

- Coxhead, Averil: "A New Academic Word List". *TESOL Quarterly*, 34(2), 2000: 213-238.
- Coxhead, Averil: "Acquiring Academic and Disciplinary Vocabulary". *The Routledge Handbook of English for Academic Purposes,* 2016: 177-190.
- Henriksen, Birgit og Jakobsen, Anne Sofie: "Akademiske ord og fraser – et nordisk projekt". *Sprogforum - Tidsskrift for sprogpædagogik*, nr. 56, maj 2013: 49-55
- Johansson, Victoria: "Lexical diversity and lexical density in speech and writing: a developmental perspective". Lund Universitet, *Working Papers* 53, 2008: 61-79.
- Nagy, William og Townsend, Dianna: "Words as Tools: Learning Academic Vocabulary as Language Acquisition". *Reading Research Quarterly*, 47(1), 2012: 91-108.
- Nation, I.S.P.: "Learning Vocabulary in Another Language". 2nd ed., Cambridge University Press, 2013: 289-321.
- Snow, C.E. og Uccelli, P. (2009): "The Challenge of Academic Language". I Olson, D.R. og N. Torrance (Eds.), *The Cambridge Handbook of Literacy*, 2009: 112-133. Cambridge: Cambridge University Press.

Andet anvendt materiale:

- Dansk Sprognævns rapport: *Dansk sprogs status 2012*, s. 10-12: https://dsn.dk/vi-arbejder-ogsa-med/sprogpolitik/sprogpolitikker-1/dansk-sprogs-status-2012/DSN_sprogstatus2012.pdf, set d. 20-10-2017

- ”Engelsk A, STX, Vejledning 2017”: https://uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017, set d. 24-10-2017
- PhraseContexts hjemmeside: http://www.hjkm.dk/phrasecontext/, set d. 24-10-2017, til estimering af pris for korpusværktøj.
- SketchEngines hjemmeside: https://www.sketchengine.co.uk/, set d. 24-10-2017, til estimering af pris for korpusværktøj.

13. Bilag

13.1 Tabel 1: Samlet oversigt for L1- og L2-datasæt

(alle tal angivet i procent)

L1-essays	**AWL-ord**	**Substantiver**	**Adjektiver**	**Præpositioner**	**Leksikalsk densitet**
Tekst 1	8	32,5	8	16	54,5
Tekst 2	4	25	7,5	14	49,5
Tekst 3	3,5	21	6	12	46,5
Tekst 4	3,5	29,5	6	13	52,5
Tekst 5	2,5	22	6	12,5	50
Tekst 6	5	29	13	14,5	57
I alt	**4,5**	**26,5**	**8**	**13,5**	**51,5**
-	-	-	-	-	-
L2-essays	**AWL-ord**	**Substantiver**	**Adjektiver**	**Præpositioner**	**Leksikalsk densitet**
Tekst A	4	25,5	5	13	46,5
Tekst B	3	23	7	14	45,5
Tekst C	3	22,5	6,5	12,5	46
Tekst D	3,5	23	4,5	14	44,5
Tekst E	4	27	6,5	13,5	46
Tekst F	2	23,5	6	13	45
I alt	**3,5**	**24**	**6**	**13,5**	**45,6**

13.2 Tabel 2: Hyppighedstabel for AWL-ord (L1-datasæt)

Nr.	**Forekommende ord fra AWL i L1-essays (127 forskellige)**	**Hyppighed**
1	computer	18
2	author	16
3	symbol	7
4	final	6
5	consequence	6
6	bond	5
7	interact	5
8	text	5
9	respond	5
10	illustrate	5
11	adult	5
12	secure	5
13	reinforce	4
14	justification	4
15	analyse	4
16	device	4
17	contrast	4
18	relevant	4
19	focus	4
20	rational	4
21	conflict	4
22	technique	3
23	negative	3
24	emphasize	3
25	demonstrate	3
26	logical	3
27	create	3

28	obvious	3
29	constant	2
30	quote	2
31	comment	2
32	theory	2
33	experts	2
34	circumstance	2
35	ultimate	2
36	brief	2
37	role	2
38	environment	2
39	seek	2
40	motivation	2
41	approach	2
42	aware	2
43	challenge	2
44	specific	1
45	despite	1
46	mutual	1
47	precisely	1
48	perception	1
49	overall	1
50	interpret	1
51	notion	1
52	participation	1
53	response	1
54	institution	1
55	economic	1
56	devotion	1
57	require	1
58	involvement	1
59	ignore	1

60	context	1
61	solely	1
62	negate	1
63	evidence	1
64	conclude	1
65	achievement	1
66	consistent	1
67	justify	1
68	reveal	1
69	ambiguities	1
70	sustained	1
71	violate	1
72	previous	1
73	aspect	1
74	impact	1
75	intelligent	1
76	facilitate	1
77	correspond	1
78	analyse	1
79	devote	1
80	task	1
81	affect	1
82	communicate	1
83	transport	1
84	obtain	1
85	shift	1
86	occur	1
87	series	1
88	whereas	1
89	internal	1
90	parallel	1
91	subsequent	1

92	link	1
93	intense	1
94	income	1
95	access	1
96	explicit	1
97	major	1
98	theme	1
99	thesis	1
100	energy	1
101	issue	1
102	deny	1
103	precede	1
104	aid	1
105	outcome	1
106	paradigm	1
107	apparent	1
108	status	1
109	contribute	1
110	community	1
111	volunteer	1
112	initial	1
113	physical	1
114	ambiguous	1
115	traditional	1
116	concept	1
117	version	1
118	abstract	1
119	crucial	1
120	section	1
121	confirm	1
122	find	1
123	somewhat	1

124	accurate	1
125	paragraph	1
126	beneficiary	1
127	strategy	1

13.3 Tabel 3: Hyppighedstabel for AWL-ord (L2-datasæt)

Nr.	**Forekommende ord fra AWL i L2-essays (70 forskellige)**	**Hyppighed**
1	text	27
2	sex	14
3	adult	14
4	symbol	13
5	theme	12
6	indicate	8
7	stability	7
8	focus	6
9	tense	6
10	tape	6
11	published	5
12	normal	5
13	author	5
14	media	4
15	perspective	4
16	significant	4
17	minor	3
18	quote	3
19	communicate	3
20	interpret	3
21	somewhat	3
22	furthermore	3
23	style	2

24	period	2
25	reliable	2
26	restricted	2
27	role	2
28	topic	2
29	contrast	2
30	react	2
31	maintain	2
32	conflict	2
33	challenge	2
34	aware	2
35	task	1
36	contrary	1
37	proceed	1
38	unpredictable	1
39	identity	1
40	instructor	1
41	obvious	1
42	concentrate	1
43	consist	1
44	evolve	1
45	odd	1
46	inappropriate	1
47	conclusion	1
48	route	1
49	technique	1
50	circumstance	1
51	location	1
52	process	1
53	vary	1
54	version	1
55	paragraph	1

56	job	1
57	enable	1
58	presumably	1
59	constant	1
60	previously	1
61	find	1
62	summary	1
63	thereby	1
64	section	1
65	finally	1
66	transform	1
67	occur	1
68	comment	1
69	aspect	1
70	parallel	1

13.4 Hyppighedstabeller for AWL-ord i de enkelte tekster (L1)

Tekst 1: **(78/986 – 50 forsk.)**	**Hyppighed**
respond	5
bond	5
justification	4
relevant	4
reinforce	3
analyse	3
text	2
comment	2
theory	2
experts	2
circumstance	2
brief	2
technique	2

device	2
contrast	2
specifically	1
despite	1
constant	1
mutual	1
precisely	1
perception	1
interact	1
overall	1
interpret	1
quote	1
author	1
notion	1
participation	1
response	1
institution	1
economic	1
devotion	1
require	1
ultimate	1
involvement	1
ignore	1
role	1
context	1
focus	1
solely	1
negate	1
evidence	1
illustrate	1
conclude	1
achievement	1

consistent	1
justify	1
reveal	1
ambiguities	1
sustained	1

Tekst 2: **(33/841 – 16 forsk.)**	**Hyppighed**
author	7
adult	5
consequence	5
negative	3
final	3
violate	1
emphasize	1
previous	1
aspect	1
impact	1
role	1
rational	1
intelligent	1
environment	1
technique	1
device	1

Tekst 3: **(49/637 – 20 forsk.)**	Hyppighed
computer	18
illustrate	3
author	3
interact	3
text	2

rational	2
create	2
logical	2
facilitate	1
demonstrate	1
correspond	1
analyse	1
focus	1
devote	1
task	1
affect	1
communicate	1
seek	1
transport	1
final	1

Tekst 4: **(37/1026 – 28 forsk.)**	**Hyppighed**
secure	5
conflict	3
emphasize	2
author	2
focus	2
consequence	1
obtain	1
rational	1
shift	1
occur	1
series	1
whereas	1
internal	1
parallel	1

subsequent	1
link	1
constant	1
environment	1
interact	1
final	1
intense	1
income	1
access	1
explicit	1
major	1
symbol	1
theme	1
motivation	1

Tekst 5: **(19/787 – 11 forsk.)**	**Hyppighed**
symbol	6
demonstrate	2
author	2
approach	2
aware	2
thesis	1
Text	1
Seek	1
energy	1
Issue	1
Deny	1

Tekst 6: **(46/949 – 40 forsk.)**	**Hyppighed**
obvious	3

contrast	2
challenge	2
precede	1
motivation	1
Aid	1
outcome	1
classic	1
conflict	1
paradigm	1
apparent	1
status	1
contribute	1
community	1
volunteer	1
initial	1
ultimate	1
physical	1
ambiguous	1
traditional	1
concept	1
author	1
version	1
abstract	1
logical	1
crucial	1
section	1
confirm	1
Find	1
create	1
somewhat	1
reinforce	1
accurate	1

Quote	1
paragraph	1
Final	1
illustrate	1
beneficiary	1
strategy	1
device	1

13.5 Hyppighedstabeller for AWL-ord i de enkelte tekster (L2)

Tekst A: **(45/1121 – 25 forsk.)**	**Hyppighed**
Text	10
Sex	7
theme	3
minor	2
Style	2
symbol	2
Quote	2
indicate	2
published	1
Task	1
Focus	1
contrary	1
period	1
media	1
proceed	1
unpredictable	1
perspective	1
identity	1
reliable	1
restricted	1
communicate	1

Role	1
significant	1
interpret	1
instructor	1
Tekst B: **(35/1157 – 20 forsk.)**	**Hyppighed**
indicate	5
Tense	4
somewhat	3
theme	3
concentrate	2
normal	2
author	2
Adult	2
Topic	1
restricted	1
Quote	1
obvious	1
period	1
consist	1
evolve	1
Sex	1
contrast	1
Odd	1
inappropriate	1
conclusion	1

Tekst C: **(34/1144 – 14 forsk.)**	**Hyppighed**
Text	7
symbol	6
perspective	3
communicate	2
Tense	2
author	2
theme	2
Adult	2
Sex	2
published	1
Route	1
normal	1
technique	1
media	1

Tekst D: **(40/1122 – 25 forsk.)**	**Hyppighed**
Text	8
stability	7
Tape	6
significant	3
Focus	3
React	2
furthermore	2
published	1
circumstance	1
location	1
process	1
Vary	1

maintain	1
version	1
paragraph	1
Job	1
enable	1
Topic	1
presumably	1
constant	1
previously	1
Find	1
summary	1
thereby	1

Tekst E: **(36/860 – 22 forsk.)**	**Hyppighed**
Adult	8
symbol	3
Focus	2
conflict	2
challenge	2
aware	2
theme	2
author	1
published	1
reliable	1
media	1
indicate	1
furthermore	1
maintain	1
text	1
sex	1

section	1
finally	1
interpret	1
reverse	1
transform	1
contrast	1

Tekst F: **(21/907 – 15 forsk.)**	**Hyppighed**
sex	3
normal	2
symbol	2
theme	2
adult	2
published	1
media	1
occur	1
text	1
comment	1
aspect	1
parallel	1
interpret	1
minor	1
role	1