

*"Samhället har nog förändrat men vi kan ju
göra det samma"*

– En korpusbaserad studie om användningen av modala
satsadverbial hos andraspråksinlärare

Ella Pantzar
Avhandling Pro Gradu
Handledare Therese
Lindström Tiedemann
Nordiska språk
Humanistiska fakulteten
Helsingfors universitet
11 / 2021

Tiedekunta – Fakultet – Faculty Humanistinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree Programme Pohjoismaisten kielten ja kirjallisuuksien maisteriohjelma	
Opintosuunta – Studieriktning – Study Track Pohjoismaiset kielet, ruotsi toisena kotimaisena kielenä			
Tekijä – Författare – Author Ella Pantzar			
Työn nimi – Arbetets titel – Title ”Samhället har nog förändrat men vi kan ju göra det samma” En korpusbaserad studie om användningen av modala satsadverbial hos andraspråksinlärare			
Työn laji – Arbetets art Level Pro gradu -tutkielma	Aika – Datum – Month and year 11/2021	Sivumäärä– Sidoantal – Number of pages 89s + liitteet	
Tiivistelmä – Referat – Abstract			
<p>Tutkielmassa tarkastellaan modaalisten lauseadverbiaalien esiintymistä oppijanruotsissa. Tutkimuksen tavoitteena on selvittää, mitkä modaaliset lauseadverbiaalit esiintyvät kielenoppijoiden teksteissä eri taitotasolla, ja miten lauseadverbiaalien käyttö eroaa L1- ja L2-ruotsissa. Kielenoppijoiden tekstejä verrataan kahteen eri ensikielen aineistoon: oppikirjateksteihin ja blogiteksteihin. Vertailulla pyritään havaitsemaan modaalisten lauseadverbiaalien mahdollista yli- tai alikäyttöä kielenoppijoiden teksteissä. Lisäksi tutkimuksessa arvioidaan automaattisesti annotoitujen korpusten soveltuvuutta tämän tyyppiseen tutkimukseen. Tutkimus on korpuspohjainen, ja analyysin teoreettisena pohjana toimii kontrastiivinen välikielianalyysi. Tutkimusmateriaali koostuu yhteensä 25:ä Göteborgin yliopiston Språkbankenin korpuksesta. Taitotasoa tutkimuksessa mitataan <i>Eurooppalaisen kielitaidon viitekehysten</i> (CEFR) mukaisesti.</p> <p>Tutkimus osoittaa, että modaalisten lauseadverbiaalien käyttö alkaa oppijankielessä taitotasolla A2 sanan <i>kanske</i> käytön alkamisen myötä. Suuremmissa määrin modaalisten lauseadverbiaalien käyttö alkaa kuitenkin vasta taitotasolla B1. Kielitaidon kehittyessä paitsi modaalisten lauseadverbiaalien esiintyvyys, myös niiden leksikaalinen variaatio teksteissä kasvaa. <i>Kanske</i> on selvästi yleisin modaalinen lauseadverbiaali oppijanruotsissa, minkä voidaan olettaa johtuvan ennen kaikkea sanan monikäyttöisyydestä. Myös <i>ju</i> esiintyy yleisesti kielenoppijoiden teksteissä tasolta B1 lähtien, mikä selittyy sanan runsaalla esiintyvyydellä sekä oppikirjateksteissä että puhekielessä. Muista modaalista lauseadverbiaaleista erityisesti <i>verkligen</i>, <i>naturligtvis</i> ja <i>egentligen</i> esiintyvät yleisesti kielenoppijoiden teksteissä.</p> <p>Vertailu ensikielen aineistoihin osoittaa, että mahdollista modaalisten lauseadverbiaalien alikäyttöä esiintyy kielenoppijoiden teksteissä sanojen <i>väl</i> ja <i>gärna</i> osalta, kun taas sana <i>naturligtvis</i> voidaan materiaalin perusteella epäillä ylikäytettävän. Varmojen johtopäätösten tekeminen tämän tutkimuksen materiaalin pohjalta ei kuitenkaan ole mahdollista, vaan edellyttää lisää tutkimusta aiheesta, sillä mahdollinen lauseadverbiaalien yli- tai alikäyttö saattaa selittyä eri tekstityyppien välisillä eroavaisuuksilla. Tutkimuksessa osoitetaan myös automaattisen annotoinnin vaikuttavan väärillä tavalla tutkimuksen lopputuloksiin, minkä johdosta tuloksiin täytyy suhtautua kriittisesti.</p> <p>Tutkimuksessa todetaan, että tutkimusmateriaalin edustavuuteen, eri materiaalien keskinäiseen vertailukelpoisuuteen sekä automaattiseen annotointiin liittyvät ongelmat vaikuttavat merkittävästi tulosten luotettavuuteen. Tämän vuoksi varmoja johtopäätöksiä modaalisten lauseadverbiaalien käytöstä oppijanruotsissa ei tämän tutkimuksen perusteella voida tehdä, vaan tulokset tulee nähdä suuntaa antavina. Tutkimus luo kuitenkin hyvän pohjan jatkotutkimuksille aiheesta ja tarjoaa myös hyödyllistä tietoa automaattisen annotoinnin tämänhetkisistä mahdollisuuksista ja haasteista.</p>			
Avainsanat – Nyckelord – Keywords Korpuslingvistiikka, modaaliset lauseadverbiaalit, kontrastiivinen välikielianalyysi, ruotsi toisena kielenä			
Säilytyspaikka – Förvaringställe – Where deposited Keskustakampuksen kirjasto			
Muita tietoja – Övriga uppgifter – Additional information			

INNEHÅLLSFÖRTECKNING

1.	INLEDNING.....	1
1.1	Syfte.....	2
1.2	Disposition.....	3
2.	TEORETISK BAKGRUND	4
2.1	Modala satsadverbial	4
2.1.1	Allmänt om satsadverbial.....	4
2.1.2	Allmänt om modala satsadverbial.....	5
2.1.3	Modalitet	7
2.1.4	Modala satsadverbial för sannolikhet och försanthållande	8
2.1.5	Modala satsadverbial för förstärkning	9
2.1.6	Modala satsadverbial för hänvisning	9
2.1.7	Modala satsadverbial för evaluering	10
2.1.8	Modala satsadverbial som kommentar till språkhandlingen.....	10
2.2	Andraspråksinläring	11
2.2.1	Förstaspråk och andraspråk.....	11
2.2.2	Inläring och tillägnande	12
2.2.3	Formell och informell inläring.....	12
2.2.4	Inlärarspråk	13
2.2.5	Bruksbaserad lingvistik och ordinläring	13
2.3	Svenska som andraspråk	15
2.4	Common European Framework of Reference for Languages (CEFR).....	16
2.4.1	Användare på nybörjarnivå (nivåerna A1 och A2).....	17
2.4.2	Självständig språkanvändare (nivåerna B1 och B2)	18
2.4.3	Avancerad språkanvändare (nivåerna C1 och C2).....	19
3.	TIDIGARE FORSKNING	20
4.	MATERIAL OCH METOD	23
4.1	Korpusar	23
4.2	Material	25
4.2.1	L2-materialet	25
4.2.2	L1-materialet	28
4.3	Metod.....	30
4.3.1	Korpustypologi.....	30
4.3.2	Contrastive Interlanguage Analysis	33

4.3.3 Sökningarna i Korp	34
5. RESULTAT OCH ANALYS.....	36
5.1 Kvantitativ analys av resultaten.....	36
5.2 Användningen av modala satsadverbial på olika kunskapsnivåer	38
5.2.1 Modala satsadverbial på nybörjarnivå	39
5.2.2 Modala satsadverbial på självständig språkanvändarnivå.....	42
5.2.3 Modala satsadverbial på avancerad nivå.....	45
5.3 Modala satsadverbial i Bloggmix-materialet.....	49
5.4 Kvalitativ analys av resultaten.....	52
5.4.1 Över- och underanvändningen av satsadverbial i inlärmaterialet.....	52
5.4.2 Användningen av <i>kanske</i>	54
5.4.3 Användningen av <i>ju</i>	58
5.4.4 Användningen av <i>väl</i>	61
5.5 Annoteringens påverkan på resultaten.....	63
5.6 Sammanfattning.....	70
6. DISKUSSION	73
6.1 Studiens resultat	73
6.2 Studiens begränsningar.....	78
6.3 Slutsatser och förslag för framtida forskning	81
LITTERATUR	84
Korpusar.....	87
BILAGA 1. Sökningarna i Korp	90
BILAGA 2. Samtliga ord som taggats som satsadverbial i L2-materialet.....	94

Tabeller

Tabell 1. CEFR-skalan.	17
Tabell 2. Studiens material.....	25
Tabell 3. Numeriska data om inlärmaterialet.....	26
Tabell 4. Antalet texter på olika CEFR-nivåer i inlärmaterialet.	26
Tabell 5. Antalet tokens på olika CEFR-nivåer i inlärmaterialet.....	27
Tabell 6. Numeriska data om läroboksmaterialet.	28
Tabell 7. De fem mest frekventa modala satsadverbialen i hela inlärmaterialet.	36
Tabell 8. De fem mest frekventa modala satsadverbialen i inlärmaterialet, taggat som satsadverbial vs alla träffar.....	37
Tabell 9. Frekvenserna av tokens som annoterats som satsadverbial i inlärmaterialet på olika kunskapsnivåer.	38
Tabell 10. De fem mest frekventa modala satsadverbialen på nivå A1 i COCTAILL. ..	40
Tabell 11. De fem mest frekventa modala satsadverbialen på nivå A2 i COCTAILL. ..	41
Tabell 12. De fem mest frekventa modala satsadverbialen på nivå B1 i inlärtexter....	42
Tabell 13. De fem mest frekventa modala satsadverbialen på nivå B1 i COCTAILL. ..	43
Tabell 14. De fem mest frekventa modala satsadverbialen på nivå B2 i inlärtexter....	44
Tabell 15. De fem mest frekventa modala satsadverbialen på nivå B2 i COCTAILL. ..	44
Tabell 16. De fem mest frekventa modala satsadverbialen på nivå C1 i inlärtexter....	45
Tabell 17. De fem mest frekventa modala satsadverbialen på nivå C1 i COCTAILL. ..	47
Tabell 18. De mest frekventa modala satsadverbialen på nivå C2 i inlärtexter.....	48
Tabell 19. De fem mest frekventa modala satsadverbialen i Bloggmix-materialet.	49
Tabell 20. Relativa frekvenserna av de mest frekventa modala satsadverbialen i inlärmaterialet och bloggmaterialet i studiens tre delmaterial.	50
Tabell 21. Sammanfattande jämförelse av användningen av modala satsadverbial i L1- och L2-svenska.....	51
Tabell 22. De relativa frekvenserna av modala satsadverbial i Bloggmix-materialet och i inlärtexter på olika kunskapsnivåer.....	51
Tabell 23. Den procentuella andelen av informanter som använder orden <i>kanske</i> , <i>ju</i> , <i>verkligen</i> , <i>egentligen</i> , <i>naturligtvis</i> och <i>inte</i> minst en gång i sin text.....	55
Tabell 24. Modala satsadverbials dependensrelationer i inlärmaterialet och i Talbanken.	65

Tabell 25. De fem mest frekventa modala satsadverbialen i hela inläromaterialet efter en manuell annotering.....	68
Tabell 26. De fem mest frekventa modala satsadverbialen på olika kunskapsnivåer i inläromaterialet.	74
Tabell 27. Frekvenserna av de fem mest frekventa modala satsadverbialen i inlärtexter på nivå C1.	80

Figurer

Figur 1. Den relativa frekvensen av tokens som annoterats som satsadverbial på olika kunskapsnivåer i inläromaterialet.	38
Figur 2. Andelen av informanter som använder orden <i>kanske, ju, verkligen, egentligen, naturligtvis</i> och <i>inte</i> i sin text samt andelen av inlärtexter som innehåller de här orden.	55

1. INLEDNING

Det händer att jag får den skämtsamma frågan om mitt vackraste svenska ord. Många brukar föreslå en vacker företeelse, *junimorgon*, ett ord med vackra ljud, *tingeltangel*, eller gärna ett ord med båda egenskaperna, *malmklang*. Jag svarar *ju*. (Josephson, 2005: 5.)

Så här inleder Olle Josephson sin bok *Ju – ifrågasätta självklarheter om svenskan, engelskan och alla andra språk i Sverige* (2005). Det här utlåtandet fungerar retoriskt sett som ett sätt att väcka läsarens intresse med någonting överraskande och oförväntat. För överraskande och oförväntat är det *ju* – det är förmodligen inte många som ens skulle ha tänkt på ordet *ju* eller något annat satsadverbial som svar på frågan om det vackraste svenska ordet. *Ju*, såsom alla satsadverbial är små ord och uttryck, som man kanske inte fäster så mycket uppmärksamhet vid, men som spelar en större roll i vårt dagliga språkbruk än vi kanske ens är medvetna om. Det är inte bara fråga om sådana ord och uttryck som är absolut nödvändiga för den språkliga framställningen, såsom t.ex. negationsordet *inte*, utan även sådana ord och uttryck som används för att uttrycka finare nyanser och som kan anses representera ett mera avancerat språkbruk, såsom t.ex. ordet *följaktligen*.

Några satsadverbial kan alltså ses representera det absolut mest grundläggande ordförrådet – det vill säga ord och uttryck som språkinläraren måste tillägna sig redan på ett tidigt stadium av språkinlärningsprocessen för att klara sig i de vardagligaste diskussionerna. Andra satsadverbial kan däremot användas för att uttrycka finare nyanser i ett språk och kan således berätta om en mycket högutvecklad språkförmåga. Det är just den här diversiteten som gör satsadverbial ett intressant studieobjekt speciellt ur

andraspråksinlärningsperspektiv: hur och i vad mån använder man satsadverbial i de olika stadierna av språkinlärningsprocessen när man håller på att lära sig ett andraspråk?¹

1.1 Syfte

I den här avhandlingen undersöker jag en specifik grupp av satsadverbial, modala satsadverbial, i inlärarsvenskan. Avhandlingen har två syften. Det första syftet är att utreda vilka modala satsadverbial som andraspråksinlärare använder i olika stadier av språkinlärningsprocessen, och se om det verkar förekomma över- eller underanvändning av ett eller flera modala satsadverbial hos andraspråksinlärare. Det andra syftet är att utvärdera hur användningen av automatiskt annoterade korpusar som forskningsmaterial lämpar sig för den här typen av studien. Användningen av annoterade korpusar är ett tvistefråga bland korpuslingvister – en del forskare opponerar sig användningen av annoterade korpusar i språkforskning för att färdiggjorda annoteringar, speciellt automatiskt genererade, kan vara inkonsekventa eller felaktiga, vilket naturligtvis kan ha en förvirrande påverkan på en studies resultat (McEnery & Hardie, 2012: 14). Den automatiska annoteringen har dock också sina obestridliga fördelar, och därför vill jag bidra till den pågående diskussionen om användningen av automatiskt annoterade korpusar som forskningsmaterial genom att under studiens gång diskutera och överväga de för- och nackdelar som finns med det här materialvalet.

Utifrån de här två syftena har jag formulerat studiens forskningsfrågor på följande sätt:

1. Vilka modala satsadverbial använder andraspråksinlärare i olika stadier av språkinlärningsprocessen?
2. Finns det över- eller underanvändning av modala satsadverbial i inlärarspråket?
3. Hur påverkar den automatiska annoteringen studiens resultat?

Forskningsmaterialet består av tre delmaterial. Det första materialet, studiens L2-material, består av tre inlärarkorpusar som innehåller texter skrivna av andraspråksinlärare. Inlärarmaterialet jämförs med två typer av L1-material: ett

¹ I den här avhandlingen står begreppet *andraspråk* (L2) för vilket som helst språk som lärs in efter förstaspråket (L1). Användningen av de här begreppen motiveras och diskuteras vidare i avsnitt 2.2.1.

läroboksmaterial, och ett bloggmaterial. Läroboksmaterialet innehåller texter från läroböcker i svenska som andraspråk, och det används för att se vilka modala satsadverbial som andraspråksinlärare överhuvudtaget exponeras för under olika faser av språkinlärningsprocessen. Bloggmaterialet består av texter från svenska bloggar, och den används som ett oavhängigt förstaspråksmaterial för att se hur resultaten förhåller sig till den naturliga språkanvändningen hos modersmålstalare. Metodologiskt sett är undersökningen korpusbaserad, och analysen baserar sig på Sylviane Grangers *contrastive interlanguage analysis* (hädanefter i avhandlingen CIA) som i andraspråksforskningen kan användas bl.a. för att undersöka om andraspråksinlärare tenderar över- eller underanvända något lingvistiskt element, t.ex. ord, fras eller struktur (Granger, 2002: 12). Resultaten analyseras framför allt kvantitativt genom att utnyttja deskriptiv statistik, men den kvantitativa analysen kompletteras även av en kort kvalitativ översikt över resultaten.

1.2 Disposition

Avhandlingen består av sex kapitel. I kapitel 2 ges en översikt av de centrala teoretiska begrepp som berör avhandlingens teman, och i kapitel 3 presenteras tidigare forskning om ämnet. Undersökningens material och metod framställs i kapitel 4. I kapitel 5 presenterar jag studiens resultat och analys, vilka sedan diskuteras i kapitel 6. I kapitel 6 behandlar jag även studiens begränsningar samt ger förslag för framtida forskning.

2. TEORETISK BAKGRUND

I det här kapitlet presenterar jag avhandlingens teoretiska bakgrund. Först i avsnitt 2.1 ges en allmän introduktion till satsadverbial, och sedan behandlar jag närmare den specifika typen av satsadverbial som ligger i studiens fokus, dvs. modala satsadverbial, samt diskuterar kort begreppet *modalitet*. Efter det, i avsnitt 2.2, förklarar jag några för den här studien centrala teman och begrepp inom andraspråksinlärning och -forskning, och i avsnitt 2.3 diskuteras svenskans ställning som andraspråk i världen, vilket visar den samhälleliga ramen som studien äger rum i. Till slut, i avsnitt 2.4, presenterar jag i korthet den så kallade CEFR-skalan som används som referensram för språkfärdigheter i den här avhandlingen.

2.1 Modala satsadverbial

I det här avsnittet presenteras ett grundläggande element för den här avhandlingen, nämligen satsadverbial. Först ger jag en kort introduktion till satsadverbial som syntaktisk kategori, och efter det riktas fokuset på modala satsadverbial och modalitet. Teorin baserar sig på Svenska Akademiens Grammatik (SAG) av Teleman m.fl. (1999a; 1999b). Exempelmeningarna är hämtade från Språkbankens PAROLE-korpus om inte annat aviserats.

2.1.1 Allmänt om satsadverbial

Enligt Teleman m.fl. (1999a: 150) är adverbial en satsdel som framför allt fungerar som en bestämning till verb, adjektiv och adverb. Vad som skiljer satsadverbial från andra adverbial är att de modifierar hela satsen, medan t.ex. tids-, rums- och sättsadverbial modifierar enstaka fraser (Svensk ordbok, 2021). Enligt Teleman m.fl. (1999b: 84) kan talaren använda satsadverbial för att uttrycka sin bedömning av det som sägs: med hjälp av ett satsadverbial kan hen bl.a. negra satsens innehåll, ange hur önskvärt det är, ta

ställning till satsens sanningshalt eller uttrycka hur satsen förhåller sig till kontexten den står i – är det t.ex. någonting som kunde förväntas, eller strider det mot det man förväntade sig.

Ett typiskt satsadverbial består av ett adverb utan bestämmingar (t.ex. *förhoppningsvis*, *åtminstone*, *inte*), men även andra fraser kan fungera som satsadverbial. Några vanliga satsadverbial utgörs av adjektivfraser med huvudordet i neutrumform (t.ex. *säkert*, *eventuellt*, *rätt och slätt*), participfraser (t.ex. *glädjande nog*, *ärligt talat*), prepositionsfraser (t.ex. *i alla fall*, *efter vad det verkar*, *för all del*) och bisatser (t.ex. *bara så du vet*, *när allt kommer omkring*). (Teleman m.fl., 1999b: 84, 86.) Komparativa subjunktionsfraser (t.ex. *som sagt*, *som väntat*) och infinitivfraser (t.ex. *av allt att döma*, *så att säga*) kan också fungera som satsadverbial, men de är ofta mer eller mindre lexikaliserade, såsom är också nominalfraser och huvudsatser som i roll av satsadverbial oftast är svordomar och eder (t.ex. *djävlar i mig*, *banne mig*) (ibid.: 86–87).

Satsadverbial kan förekomma såväl i huvud- och bisatser som i vissa satsekvivalenta fraser, såsom t.ex. infinitivfraser. I finita satser placeras satsadverbial oftast i satsens mittfält, men det är även möjligt att placera satsadverbial i satsens fundament eller slutfält. (Teleman m.fl., 1999b: 88–89.)

- (1) Pompeiji är **ju** ett av den moderna världens antika underverk.
- (2) **Tyvärr** är industrisamhällets ekonomiska system inte avpassat för minskande folkmängd.
- (3) Jag tyckte synd om honom, det gjorde jag **verkligen**.

Satsadverbial kan delas in i olika undergrupper enligt deras funktion. I SAG delas satsadverbial in i fyra huvudgrupper som kallas för *modala*, *konjunktionella*, *fokuserande* och *negerande satsadverbial* (Teleman m.fl., 1999b: 84). Eftersom den här avhandlingen fokuserar på modala satsadverbial, presenteras den här underkategorin närmare i följande underkapitel.

2.1.2 Allmänt om modala satsadverbial

Modala satsadverbial är en typ av satsadverbial som enligt Hultman (2003: 276) uttrycker hur talaren förhåller sig till satsinnehållet. Teleman m.fl. (1999b: 84) utgår från en definition att modala satsadverbial särskilt anger talarens bedömning av hur sannolikt satsinnehållet är, eller hur talaren värderar det som sägs, om det t.ex. är önskvärt eller

obehagligt. Med modala satsadverbial kan talaren även ange vilken reaktion hen väntar från lyssnaren eller ”röja vilka föreställningar talaren har om lyssnarens attityd till satsinnehållet” (Teleman m.fl., 1999b: 96).

Modala satsadverbial förekommer i huvudsatser och bisatser, främst i hävdade eller presupponerade satser eller fraser. De kan även elliptiskt utgöra egna yttranden (ex. 6), något som inte är lika vanligt hos andra typer av satsadverbial. (Teleman m.fl., 1999b: 88–89.)

- (4) Arkebakterierna är **kanske** den form av liv som har funnits längst på jorden.
- (5) Charles beställer en rom och Coca-Cola, han föreställer sig att det **antagligen** är sånt man dricker på Bermuda.
- (6) – Är den stulen? – *{Troligen/Lyckligtvis}* (Teleman m.fl., 1999b: 89). (Min emfas).

Modala satsadverbial står vanligen i satsens mittfält. Om det finns fler än ett modalt satsadverbial i mittfältet, placeras de obetonade modala satsadverbialen (*ju, väl, nog, då*) först och andra modala satsadverbial följer i en vedertagen ordning enligt följande schema av SAG:

X	Språkhandlingskommentar	Hänvisning	Förstärkning	Evaluering	Sannolikhet	Y
My har	ärligt talat	enligt Per	faktiskt	tyvärr	antagligen	gått.
De vill	apropå det	i princip	minsann	turligt nog	kanske	gå.

SCHEMA 2. De modala satsadverbialens ordning i mittfältet.

Bild 1. De modala satsadverbialens inbördes ordning i mittfältet (Teleman m.fl., 1999b: 98, inklusive bilden).

I satser där det finns flera olika typer av satsadverbial i mittfältet placeras det negerande satsadverbialet oftast sist, men några modala satsadverbial kan ändå komma efter det negerande satsadverbialet (ex. 7). De står då tillsammans på det negerande satsadverbialets plats. (Teleman m.fl., 1999b: 92.) Detsamma gäller för satsadverbial som kan ses som både modala och negerande. Också de placeras vanligtvis sist i mittfältet, såsom övriga negerande satsadverbial (ex. 8). (ibid.: 91.)

- (7) Svenskarna vill **inte gärna** gå i kyrkan; men de vill ha den kvar.
- (8) Att universitetsledningen nu handlat som den gjort kan **alltså knappast** motiveras med vetenskapliga argument.

De flesta modala satsadverbialen kan inleda deklarativa satser, dvs. placeras i satsens fundament. Bland de modala satsadverbial som inte kan stå i fundamentet är t.ex. *ju*, *verkligen*, *för all del*, *förstås*, *nödvändigtvis* och *väl*. Det är inte heller vanligt att placera sådana satsadverbial som börjar med prefixet *o-* eller är av svordomskaraktär i fundamentet. (Teleman m.fl., 1999b: 99.)

(9) **Kanske** var det inte så förskräckande att solsystemet blev dubbelt så stort, eftersom det redan var ofattbart stort.

Det är även möjligt att placera modala satsadverbial i satsens slutfält. Detta gäller för de flesta modala satsadverbialen förutom de modala satsadverbial som har en negerande betydelse. I satsens slut står dessutom vanligtvis bara ett satsadverbial. (Teleman m.fl., 1999b: 100.)

(10) Erfarenheten visar att det **knappast** är en nackdel att sakna politisk bakgrund.

(11) *Erfarenheten visar att det är en nackdel att sakna politisk bakgrund **knappast**. (Min modifiering)

(12) Det händer **väl** att man tar chansen när den bjuds **så att säga**.

(13) *Det händer att man tar chansen när den bjuds **väl så att säga**. (Min modifiering)

2.1.3 Modalitet

Modala satsadverbial är nära förknippade med begreppet *modalitet*. Det råder inte enighet bland forskare om vad begreppet egentligen täcker: medan några språkforskare anser att det handlar bara om utsagans möjlighet eller nödvändighet, ser andra forskare ett större fenomen bakom begreppet (von Steutterheim, 1993: 4). Gemensamt för alla de olika definitionerna tycks dock vara att modalitet anses på något sätt beskriva talarens inställning till satsinnehållet. Den här breda definitionen används bl.a. av Lagerholm (2008: 152) och Muittari som definierar modalitet som ”en semantisk kategori som berör talarens inställning till handlingen eller till dess verklighetsstatus” (1988: 80). I svenskan uttrycks modalitet ofta med hjälp av modala satsadverbial. Av de modala betydelse som modala satsadverbial uttrycker kan enligt Teleman m.fl. (1999b: 96–97) urskiljas fem grundbetydelser som presenteras i följande underkapitel.

2.1.4 Modala satsadverbial för sannolikhet och försanthållande

Modala satsadverbial som tar ställning till satsinnehållets sannolikhet kan delas in i två grupper enligt hurdan uppfattning talaren har om satsinnehållet. Den första gruppen är satsadverbial som graderar sannolikheten hos satsen (ex. 14), och den andra gruppen består av satsadverbial som graderar satsens försanthållande (ex. 15). I det senare utgår talaren alltså från att satsinnehållet är sant. (Teleman m.fl., 1999b: 102.)

(14) Hon är **antagligen** i Colorado.

(15) Han var **uppenbarligen** mycket uppskärnad.

Bland de vanligaste satsadverbialen för sannolikhet nämner Teleman m.fl. (1999b: 104) t.ex. *antagligen*, *eventuellt*, *näppeligen*, *förmodligen*, *garanterat*, *kanske*, *knappast*, *(o)möjligen*, *nog*, *sannolikt*, *säkert*, *troligen* och *väl*. De representerar olika grader av sannolikhet: medan t.ex. *näppeligen*, *knappast* och *omöjligen* uttrycker att det är osannolikt att det som sägs är sant, uttrycker t.ex. *garanterat*, *troligen* och *säkert* att satsinnehållet är mycket sannolikt. Däremellan finns sådana neutrala satsadverbial som uttrycker att satsinnehållet kan vara sannolikt eller osannolikt, men tar inte närmare ställning till sannolikhetsgraden. Sådana satsadverbial är t.ex. *kanske*, *möjligen* och *eventuellt*. (ibid.: 103.)

(16) Studenter kan **näppeligen** betraktas som en svag grupp.

(17) Är det fint väder åker jag **garanterat** till Saltsjöbaden och badar.

(18) Vaccination mot allergi kan **kanske** bli möjlig inom en rimlig framtid.

Modala satsadverbial som uttrycker försanthållande används enligt Teleman m.fl. (1999b: 105) för att framhäva eller bekräfta satsens sannhet, t.ex. om den ifrågasätts av lyssnaren. Som exempel på satsadverbial som direkt uttrycker att satsen är sann och inte lämnar utrymme för andra tolkningar nämner Teleman m.fl. (ibid.) t.ex. *absolut*, *definitivt*, *otvivelaktigt*, *utan tvivel*, *obestridligen* och *onekligen*.

(19) Regeringens insatser har **onekligen** gett resultat.

Satsadverbial som däremot ”anger att försanthållandet är en rimlig, självklar eller nödvändig slutsats” (Teleman m.fl., 1999b: 105) är t.ex. *av allt att döma*, *rimligen*, *som det vill synas*, *tydligt*, *förstås*, *givetvis*, *ju*, *naturligtvis*, *självklart*, *visst* och *nödvändigtvis*.

(20) Sveriges största problem är **självklart** arbetslösheten.

2.1.5 Modala satsadverbial för förstärkning

Modala satsadverbial för förstärkning fungerar som effektmedel för att ge styrka till utsagan. SAG särskiljer två typer av förstärkande satsadverbial. Den första typen är satsadverbial som uttrycker förvåning eller ”en garanti för sanningen i ett påstående eller uppriktigheten i ett löfte eller ett hot – ofta gentemot vad talaren själv eller lyssnaren hade väntat sig” (Teleman m.fl., 1999b: 106). Till den här gruppen hör t.ex. många svordomar, såsom *banne mej* och *jävlar i mig*, men också mera neutrala ord såsom *minsann*, *faktiskt* och *verkligen*. Även *nog* kan fungera som förstärkande satsadverbial om det står i satsens fundament. (ibid.)

(21) Nu får du **banne mej** ligga i din egen säng.

(22) Det är **minsann** inte alltid så lätt att vara.

(23) **Nog** är det märkligt!

Till den andra gruppen av förstärkande modala satsadverbial hör ord och uttryck som anger talarens irritation men också förstärker språkhandlingen. De flesta satsadverbialen i den här gruppen är svordomar eller svordomsersättningar, t.ex. *för {fan/tusan/sjutton/allt i världen}*. (Teleman m.fl., 1999b: 107.)

(24) Det kan ju **för tusan** inte finnas något mord utan ett lik.

2.1.6 Modala satsadverbial för hänvisning

Modala satsadverbial för hänvisning används enligt Teleman m.fl. (1999b: 108) för att göra lyssnaren medveten om olika faktorer som kan påverka satsinnehållet. De särskiljer tre olika funktioner för hänvisande modala satsadverbial: satsadverbial som hänvisar till en källa (t.ex. *enligt X, som X säger, såsom tidigare nämnts*), satsadverbial som hänvisar till begränsningar i talarens perceptuella och mentala förmåga som kan påverka satsens sanningshalt (t.ex. *efter vad det verkar, efter vad jag vet, enligt min mening, så vitt jag vet*), samt satsadverbial som används för att ”hänvisa till en viss synvinkel eller ett speciellt betraktelsesätt som måste accepteras för att satsen skall vara sann” (Teleman m.fl. 1999b: 109). Till den sistnämnda gruppen hör t.ex. *egentligen, i princip, i praktiken, i realiteten, i själva verket* och *i stort sett*. (ibid.: 108–109.)

(25) En av bussarna har **enligt polisen** inte följt väjningsplikten och stannat.

(26) Han bor ensam och är **så vitt jag vet** helt harmlös.

(27) Eleverna är **i stort sett** nöjda med undervisningen.

2.1.7 Modala satsadverbial för evaluering

Evaluerande modala satsadverbial anger hur talaren värderar satsinnehållet, om det t.ex. anses vara positivt eller negativt, förväntat eller överraskande o.s.v. Satsadverbialen kan också ange hurdan reaktion som sakförhållandet väcker i talaren. (Teleman m.fl., 1999b: 110.)

- (28) Men **lyckligtvis** går det snabbt att reparera!
- (29) Jag kommer **tyvärr** inte ihåg vad han hette.
- (30) Kritik söker sig **naturligt nog** till det mest framträdande.
- (31) Där ute ser Sankt Eriksplan, **konstigt nog**, ut precis som vanligt.

2.1.8 Modala satsadverbial som kommentar till språkhandlingen

Den sista kategorin för modala satsadverbial är satsadverbial som fungerar som talarens kommentar till språkhandlingen. Även den här kategorin kan delas in i två undergrupper. Den första gruppen är modala satsadverbial som karakteriserar talarens egen språkhandling, och som ofta utgörs av ett sättsadverbial som bestämmer ett perfektparticip, t.ex. *kort sagt, ärligt talat, fritt översatt* och *med andra ord*. (Teleman m.fl., 1999b: 112.)

- (32) **Kort sagt** krävs det ett mer kontinentalt förhållningssätt till öl, vin och sprit.
- (33) Nattnattlivet är **milt sagt** föga utvecklat och slutar tidigt.

Den andra typen av modala satsadverbial som kommenterar språkhandlingen är satsadverbial som talaren använder för att förklara för lyssnaren vad som får hen att uttrycka det som hen säger. De här satsadverbialen förekommer ofta i form av en adverbial bisats. (Teleman m.fl., 1999b: 113.)

- (34) Det verkar lite konstigt **nu när du säger det**.

2.2 Andraspråksinläring

I det här kapitlet ger jag en översikt över de begrepp och fenomen inom andraspråksinläringen och -forskningen som är centrala med avseende på den här avhandlingen.

2.2.1 Förstaspråk och andraspråk

I språkforskning är det meningsfullt att skilja mellan *förstaspråk* och *andraspråk*. Förstaspråk står för det språk som man som barn lär sig först – alltså det språk som föräldrar eller andra vårdnadshavare talar för barnet. Andraspråk kan däremot anses stå för alla de språk som lärs in efter förstaspråket. (Abrahamsson, 2009: 13.) Ibland kan det vidare skiljas mellan *andraspråk* och *tredjespråk*, t.ex. i studier som fokuserar på förstaspråkets och andraspråkets påverkan på inläringen av ett tredje språk, men såsom Abrahamsson (2009: 14) påpekar, används begreppet andraspråk ofta som paraplyterm som täcker vilket som helst språk som inte är ens förstaspråk.

Ibland är det också viktigt att göra en distinktion mellan begreppen *andraspråk* och *främmandespråk*. Som främmandespråk räknas då sådana språk som en individ lär sig utanför språkets naturliga omgivning, medan ett andraspråk lärs in i en miljö där språket i fråga används som huvudsakligt kommunikationsspråk (Abrahamsson, 2009: 14). Hammarberg (2004: 26) konstaterar att när man talar om svenska som andraspråk i Sverige, kan begreppet ofta tolkas i den trängre meningen. Övriga språk, t.ex. engelska, kan i Sverige kallas för främmandespråk, eftersom de lärs in utanför den naturliga språkmiljön.

I den här avhandlingen används begreppet *andraspråk* (L2), för att beskriva vilket som helst språk som har lärts in efter förstaspråket (L1), oberoende av om inläringen kan anses ha skett i språkets naturliga miljö eller inte. Det här valet är motiverat för att såsom Abrahamsson (2009: 15) påpekar, är distinktionen mellan andraspråk och främmandespråk inte nödvändigtvis relevant att göra om studiens syfte inte är att jämföra just de här två specifika varianterna.

2.2.2 Inläring och tillägnande

En annan viktig distinktion inom andraspråksforskningen är den mellan *inläring* och *tillägnande*. Abrahamsson (2009: 15–16) definierar inläringen (eng. *learning*) som en medveten och formfokuserad process som syftar på att man lär sig att kontrollera och korrigera sitt eget språkbruk. Med tillägnande (eng. *acquisition*) menar han däremot en mer automatisk och naturlig process som leder till att man lär sig att verkligen producera och förstå språk. Till skillnad från inläringen, som oftast äger rum i klassrumsmiljö, sker tillägnandet i naturlig interaktion med målspråkstalare. (Abrahamsson, 2009: 15.) Utgående från de här definitionerna beskriver Abrahamsson (2009: 15–16) den språkkunskap som skaffas via inläring som *explicit*, medan tillägnandet leder till *implicit språkkunskap*.

Abrahamsson (2009: 16) konstaterar att man på svenska föredrar begreppet *inläring* som en paraplyterm för såväl den explicita språkinläringen som det implicita tillägnandet. Därför används termen *andraspråksinläring* som ett omfattande begrepp för båda fenomenen också i den här avhandlingen.

2.2.3 Formell och informell inläring

En distinktion som allt oftare lyftas fram är den mellan *formell* och *informell inläring*. Med begreppet *formell inläring* hänvisas enligt Abrahamsson (2009: 16) till den traditionella, undervisningsbaserade inläringen som äger rum i klassrumsmiljö, medan den informella inläringen sker utanför klassrummet, t.ex. i interaktion med målspråkstalare. På sätt och vis kan den här indelningen alltså anses sammanfalla med distinktionen mellan inläring och tillägnande – i princip kan den formella vägen sägas leda till inläring, medan den informella vägen leder till tillägnande.

I praktiken utesluter den formella och den informella inläringen dock inte varandra. Abrahamsson (2009: 16) anmärker att t.ex. inläringen av ett främmandespråk, som traditionellt sker formellt i främmandespråksundervisningen, kan likaväl ske med hjälp av TV, musik och litteratur, dvs. via den informella vägen, medan andraspråksinläringen som däremot kan ske nästan av sig själv via den språkliga exponeringen i en omgivning där språket naturligt talas i, stöds ofta av t.ex. språkkurser, och den är således också en blandning av båda inläringstyperna. Även Hammarberg (2004: 26) konstaterar att

andraspråksinläringen av svenska i Sverige ofta sker via båda inläringstyperna, dock i varierande proportioner. Med tanke på informanterna i den här undersökningen är det grundat att anta att språkinlärningsprocessen har också för dem varit en blandning av informell och formell inläring.

2.2.4 Inlärarspråk

Ett viktigt begrepp med tanke på den här undersökningens metodologi är *inlärarspråk*. I slutet av 1960-talet började man betrakta inlärarens andraspråk som ett eget system som håller på att utvecklas (Hammarberg, 2004: 27). Begreppet *interimspråk* (eng. *interlanguage*), eller på svenska numera oftare *inlärarspråk*, lanserades av lingvisten Larry Selinker i 1972 i samband med den så kallade interimspråksteorin, men nuförtiden används begreppet som en teorineutral term (Abrahamsson, 2009: 42–43). Den nya termen skapades för att skilja mellan *målspråket* som ett föremål för inläringen och *inlärarspråket* som den version av målspråket som inläraren själv bygger upp och som är ”ett successivt tillväxande system” (Hammarberg, 2004: 27). Att göra en distinktion mellan inlärarspråket och målspråket är nödvändigt eftersom inlärarspråket och målspråket kan skilja sig markant från varandra. Norrby och Håkansson (2007: 69) skriver att några typiska drag för inlärarspråket är t.ex. förenklingar, övergeneraliseringar och *transfer*, dvs. inslag av andra språk. Även Abrahamsson (2009: 14, 262) beskriver inlärarspråket som en enklare, mer föränderlig och mindre stabil version av målspråket, men framhäver vikten av att inte betrakta inlärarspråket som en felaktig version av målspråket. Den centrala tanken med inlärarspråket är enligt Hammarberg (2004: 27) att det ska ses som en konstruktiv process där inläraren själv är en aktiv aktör.

2.2.5 Bruksbaserad lingvistik och ordinläring

En av de mest framträdande inriktningarna inom den nutida språkforskningen är den så kallade *bruksbaserade lingvistik* (BBL) som speciellt inom andraspråksforskningen har blivit alltmer populär. Den bruksbaserade synen på språkinläringen utgår från tanken att det finns ett kontinuerligt samspel mellan *språkssystem*, dvs. den språkliga kunskapen som finns lagrad i hjärnan, och *språkbruk*, alltså de kommunikativa aktiviteterna där språket används. De här två delarna av språket har traditionellt hållits isär, men i BBL integreras

de till varandra: den grundläggande principen i BBL är att det omgivande språkbruket påverkar språksystemet. (Lyngfelt & Prentice, 2020: 1–2.)

Ordförråd spelar en avgörande roll i andraspråksinlärarens språkliga kompetens – ord behövs såväl för att kunna lyssna och läsa, som för att själv kunna tala och skriva. Uppbyggandet av ett ordförråd kan beskrivas som en ändlös process, och även för mer avancerade andraspråksinlärare brukar det ofullständiga ordförrådet vara den faktorn som mest begränsar deras språkliga framställning. (Enström, 2010: 11, 21.) Kunskap om ord hör centralt till språksystem, men traditionellt har det ansetts omfatta bara uttryckens form, betydelse och funktion, i praktiken alltså kunskapen om vad som är grammatiskt och vad som är ogrammatiskt. Kunskapen om vad som däremot är vanligt och vad som är ovanligt inom det grammatiska har setts ligga utanför grammatiken, och således också utanför det språkliga systemet, men inom BBL anses även brukligheten hos språkliga uttryck vara en integrerad del av ens språksystem. (Lyngfelt & Prentice, 2020: 2.) I den språkliga kunskapen om satsadverbial ingår således inte bara förmågan att använda dem grammatiskt korrekt, utan också kännedom om vilka stilistiska nyanser de meddelar, t.ex. att satsadverbial *ju* kan i vissa kontexter ha en mycket talspråklig klang.

Att lära sig nya ord förutsätter att det finns *input* av något slag, alltså att man utsätts för okända ord i skriftliga eller muntliga sammanhang (Enström, 2010: 23). Att inläraren möter ett nytt uttryck upprepade gånger och så småningom börjar också själv använda det kan kallas för *befästning* (eng. *entrenchment*) (Lyngfelt & Prentice, 2020: 5). Läroböcker har naturligtvis en central roll som kanal för språklig input, trots att såsom Enström (2010: 29) påpekar, är det bara en bråkdel av ord som lärs in i undervisningen. Undervisningens roll bör ändå inte underskattas – enligt Enström (2010: 30) är ett av undervisningens primära syften att hjälpa inläraren att organisera mångfalden i ordförrådet, t.ex. att få hen att uppmärksamma vilken relation ett ord har till andra ord, eller vilken stilnivå eller värdeladdning ett ord kan representera, vilket för sin del utvecklar inlärarens språksystem och således påverkar också språkbruket. Även Lyngfelt och Prentice (2020: 11) konstaterar att undervisningen faktiskt kan ha en avgörande roll i hur andraspråksinlärarens ordförråd blir till: om t.ex. ett ord förekommer mer frekvent i undervisningen än det förekommer i det allmänt konventionaliserade språkbruket utanför klassrummet, kan det formas en systematisk förvrängning i inlärarens version av målspråket. Därför anser jag att det är viktigt i den här avhandlingen att jämföra

inlärartexter också med lärobokstexter för att se om det av L1-svenskan möjligen avvikande språkbruket kan förklaras med läroböckernas inverkan.

2.3 Svenska som andraspråk

Svenskan är huvudspråk i Sverige och ett nationalspråk i Finland. Totalt har språket över 10 miljoner talare vilket gör det på en global nivå stort språk. Förutom modersmålstalare har svenskan även ett stort antal andraspråkstalare. I Sverige undervisas svenska som andraspråk för elever med ett annat modersmål än svenska, och den är också integreringsspråket för invandrarna i Sverige. (Svenska institutet, 2020.)

De som lär sig svenska som andraspråk i Sverige är inte någon marginell grupp. År 2019 beräknades det att över 25 % av Sveriges befolkning har en utländsk bakgrund, dvs. är själv födda utanför Sverige eller har två utrikesfödda föräldrar (Migrationsinfo.se, 2020). I dagens samhälle är det dock inte lätt att dra en gräns mellan förstaspråks- och andraspråkstalare, såsom poängteras även av Hyltenstam och Lindberg (2004: 15). Det finns t.ex. allt fler barn som är födda i Sverige och som har växt upp i det svenska samhället, men vars föräldrar har flyttat till Sverige från ett annat land och det primära språket i hemmet har således varit något annat än svenska. Frågan är då om de kan räknas som andraspråksinlärare oavsett vilket språk som är starkare för dem, eller vilket språk som de själv betraktar som deras förstaspråk: språket som enbart talas hemma med familjen eller språket som de använder i alla andra kontexter utanför hemmet.

Ett stort antal andraspråkstalare har svenskan också i Finland där svenskan är ett officiellt språk som talas som modersmål av ca 5,2 % av befolkningen (Statistikcentralen, 2020: 8). Tack vare den officiella ställningen är svenskan ett obligatoriskt ämne i den finska grundskolan också för dem som inte talar svenska som modersmål. Undervisningen i svenska som andraspråk² börjar senast i sjätte klass i den finskspråkiga grundutbildningen

² Noga taget kallas den svenskan som i Finland undervisas för studenter som deltar i modersmålsundervisningen i finska eller samiska för svenska som ”andra inhemska språk”, medan undervisning i svenska som andraspråk erbjuds för studenter som inte talar finska, svenska eller samiska som modersmål eller som i övrigt har en flerspråkig bakgrund (Utbildningsstyrelsen, u/å). Enligt den definitionen som i den här avhandlingen står för andraspråk kan även undervisningen i svenska som andra inhemska språk räknas som andraspråksundervisning.

(Språkläraryförbundet i Finland, u/å). I princip kan det alltså hävdas att alla finländare som har gått i skolan i Finland efter grundskolereformen³ och inte talar svenska som modersmål talar svenska som andraspråk. Det är även möjligt för invandrare att välja att studera svenska som integreringsspråk i Finland (Helsingfors stad, u/å).

Förutom Sverige och Finland finns det människor som talar svenska som andraspråk också i andra delar av världen, t.ex. i USA där språket lever mer eller mindre kvar hos ättlingar till svenska emigranter. Dessutom bedrivs undervisning i svenska vid över 200 universiteterna runt om i världen. (Svenska institutet, 2020.)

Forskningen kring svenska som andraspråk fick sitt startskott till följd av den stora arbetskraftsinvandringen till Sverige på 1970-talet (Hammarberg, 2004: 25). Sedan dess har forskningen i svenska som andraspråk varit aktiv speciellt i Sverige och i Finland, och trots den relativt korta historien har svenskan redan blivit ett av de mest dokumenterade andraspråken i världen (Norrby & Håkansson, 2007: 77). Mer förståelse och kunskap om svenska som andraspråk behövs dock fortgående, inte minst på grund av den ständigt ökande invandringen. Forskning i svenska som andraspråk ger naturligtvis även kunskap och förståelse om andraspråksinlärning och andraspråksinlärare på en mer generell nivå, vilket bidrar således hela den moderna språkforskningen.

2.4 Common European Framework of Reference for Languages (CEFR)

The Common European Framework of Reference for Languages (sv. Gemensam Europeisk Referensram för Språk), som hädanefter i avhandlingen kallas med den internationella akronymen CEFR, är Europarådets referensram för värderingen av språkfärdigheter. Referensramen skapades för att erbjuda en enhetlig, tydlig och objektiv beskrivning av mål, innehåll och metoder som kan utnyttjas i språkutbildning, t.ex. i undervisning, kurs- och läroplaner samt i utarbetning av läromedel i hela Europa. Referensramen kan inte bara vara till nytta för dem som arbetar inom språkutbildningen,

³ Ett obligatoriskt skolämne blev svenskan i finska skolor först i samband med grundskolereformen år 1972–1977, men det var dock vanligt att finskspråkiga elever studerade svenska i skolan också dessförinnan (Himberg, 2013).

utan också de som håller på att lära sig ett nytt språk kan utnyttja den t.ex. i självvärdering och som ett verktyg för att sätta upp mål. (Skolverket, 2009: 1, 6.)

Europarådet rekommenderar användningen av CEFR i språkutbildningen som ett sätt att ”främja den demokratiska medborgarskapen, den sociala sammanhållningen och den interkulturella dialogen” (Europarådet, 2020: 27, min översättning). Den är dock inte ett helt oproblematiskt sätt att utvärdera språkfärdigheter, såsom har poängterats av bl.a. Figueras (2012) och Harsch (2018), men eftersom den ändå skapar en tämligen användbar och tydlig beskrivning för språkkunskaper och dessutom är allmänt känd och godtagen runtom i världen, används den som referensram också i den här avhandlingen.

Referensramen består av sex kunskapsnivåer som ofta delas in i tre breda kategorier (se tabell 1).⁴ Eftersom den sexdelade uppdelningen används också i studiens material, presenteras alla de sex delnivåerna i följande avsnitt i korthet och med tanke på de färdigheter som är relevanta för den här undersökningen. En mer detaljerad beskrivning av de olika nivåerna är tillgänglig t.ex. i boken *Gemensam Europeisk Referensram för Språk: Lärande, Undervisning och Bedömning* (Skolverket, 2009).

<i>Användare på nybörjarnivå</i>	A1
	A2
<i>Självständig språkanvändare</i>	B1
	B2
<i>Avancerad språkanvändare</i>	C1
	C2

Tabell 1. CEFR-skalan (Skolverket 2009: 23).

2.4.1 Användare på nybörjarnivå (nivåerna A1 och A2)

På nivåerna A1 och A2 kallas den lärande för *användare på nybörjarnivå*. På den allra lägsta nivån, A1, beskrivs inläraren kunna t.ex. berätta om sig själv och besvara frågor angående personliga förhållanden. Ett samtal med andra språkanvändare är möjligt men förutsätter att samtalspartnern talar långsamt och tydligt. (Skolverket, 2009: 24.) För

⁴I den nya uppdateringen av CEFR indelas de här sex nivåerna ännu i så kallade plusnivåer (t.ex. A2 och A2+, eller A2 och A2.1), som beskrivs av Europarådet på följande sätt: ”Plus levels represent a very strong competence at a level that does not yet reach the minimum standard for the next criterion level. Generally, features of the level above are starting to appear” (Europarådet, 2020: 38). I den här avhandlingen används den ursprungliga sexdelade skalan för att den skalan används också i studiens material.

ordförrådets del beskrivs inläraren på nivå A1 ha ett grundläggande ordförråd som består av välkända och vardagliga ord och fraser som möter konkreta behov (ibid.: 24, 109). Grammatisk kompetens ligger i sin linda, och inläraren visar bara en begränsad behärskning av några få enklaste grammatiska strukturer (ibid.: 111).

På nivå A2 kan inläraren kommunicera i enkla sammanhang och kan med enkla ord berätta t.ex. om sin egen bakgrund och sin närmaste omgivning. Hen förstår vanliga uttryck som hänför sig vardagliga situationer, t.ex. inköp, sysselsättning och lokal geografi. (Skolverket, 2009: 24.) Ordförrådet på nivå A2 är tillräckligt stort för att inläraren kan uttrycka sig i grundläggande, rutinmässiga vardagshandlingar (ibid.: 109). För grammatikens del beskrivs inläraren på nivå A2 kunna använda några enkla strukturer korrekt, men grundläggande grammatiska fel förekommer fortfarande systematiskt (ibid.: 111).

2.4.2 Självständig språkanvändare (nivåerna B1 och B2)

På de mellersta kunskapsnivåerna, nivåerna B1 och B2, kallas språkinläraren för *självständig språkanvändare*. På nivå B1 förstår inläraren huvudinnehållet i det hen hör eller läser i vardagliga situationer, om språket är tydligt och standardiserat. Hen kan producera enkla men sammanhängande texter om välkända ämnen och kan beskriva erfarenheter och händelser samt förhoppningar, drömmar och framtidsplaner. På nivå B1 kan språkinläraren också redan kortfattat motivera sina planer och åsikter. (Skolverket, 2009: 24.) Ordförrådet täcker de flesta vardagliga ämnena, såsom familj, hobbyer och arbete, men det kan fortfarande vara svårt för inläraren att uttrycka sig själv i mer komplexa tankegångar (ibid.: 109). Grammatisk kompetens är tillräcklig utvecklad för att inläraren kan i förutsägbara situationer använda tämligen korrekta frekventa sats- och ordstrukturer (ibid.: 111).

Inläraren på nivå B2 kan kommunicera med modersmålstalare utan större ansträngningar för någondera parten. Hen producerar detaljerade och tydliga texter om flera ämnen och kan beskriva för- och nackdelar vid valmöjligheter. (Skolverket, 2009: 24.) Ordförrådet är stort när det gäller de flesta allmänna ämnen, och det används mångsidigt så att det t.ex. inte blir alltför många upprepningar av ett och samma ord. Den lexikala korrektheten är generellt stor, men omformuleringar, hopblandningar och felaktiga ordval kan ännu förekomma. (ibid.: 109.) Inläraren på nivå B2 har en relativt högutvecklad grammatisk

kompetens – enstaka grammatiska fel kan ännu förekomma men inte systematiskt. De eventuella felen leder inte till missförstånd och de kan ofta korrigeras i efterhand. (ibid.: 111.)

2.4.3 Avancerad språkanvändare (nivåerna C1 och C2)

Nivåerna C1 och C2 representerar de högsta möjliga kunskapsnivåerna på CEFR-skalan. Inläraren kallas då för *avancerad språkanvändare*. På nivå C1 kan inläraren bl.a. förstå mera komplexa och krävande texter och uttrycka sig utan några större ansträngningar spontant och flytande. Hen producerar välstrukturerade och detaljerade texter som formar sammanhängande helheter. (Skolverket, 2009: 24.) Ordförrådet täcker såväl vardagliga som idiomatiska uttryck, och inläraren kan utan besvär kringgå luckor i ordförrådet (ibid.: 109). Inläraren behärskar grammatiken i genomgående hög grad, fel uppstår bara sällan, och de är dessutom ofta svåra att upptäcka (ibid.: 111).

Inläraren på den högsta kunskapsnivån, C2, kan förstå praktiskt taget allt hen läser eller hör och kommunicerar också själv tydligt, spontant och precist också i mera komplexa situationer. Hen kan sammanfatta information och argumentera på ett sammanhängande sätt. (Skolverket, 2009: 24.) Inläraren har ett brett ordförråd som hen kan använda på korrekt och lämpligt sätt (ibid.: 109). Den grammatiska kompetensen är högutvecklad även när det gäller mer komplext språk eller situationer där uppmärksamheten är riktad på annat (ibid.: 111). Det bör påpekas att det inte ens på nivå C2 förväntas att inläraren använder språket på ett lika avancerat och flytande sätt som en modersmålstalare, trots att den nivån ibland på engelska kallas för *mastery* (Europarådet, 2020: 37).

3. TIDIGARE FORSKNING

I det här kapitlet presenterar jag tidigare forskning som tangerar den här avhandlingens teman och metodologi.

Satsadverbial i sig, speciellt de modala, verkar vara ett rätt oforskat ämne inom andraspråksforskningen i svenska, men det finns en rad studier av svenska som andraspråk där satsadverbialets placering studeras i samband med ordföljden. Speciellt ofta handlar det om det negerande satsadverbialets placering som har ur andraspråksperspektiv studerats av bl.a. Hyltenstam (1978), Bolander (1987) och Håkansson (1989). Satsadverbialets placering har även studerats ur andra perspektiv, t.ex. av Andréasson (2007) som i sin doktorsavhandling studerade satsadverbialets placering i förhållande till L1-svenskans informationsdynamik.

Modala satsadverbial omfattar även några modalpartiklar, såsom *ju*, *väl* och *nog*, trots att de noga taget inte kan ses som ordinarie satsadverbial eftersom de styrs av delvis annorlunda syntaktiska regler (Lindström, 2008: 95). Modalpartiklar i svenskan har ur kontrastivt perspektiv studerats av bl.a. Aijmer (1996; 2015; 2016). Modala satsadverbial har dessutom varit ett föremål för intresse i några mindre avhandlingar, t.ex. av Forsberg (1988) som jämförde användningen av modala satsadverbial i finlandssvenskt och sverigesvenskt tidningsspråk, och Petersson (2008) som studerade hurdana betydelser satsadverbialen *inte*, *nog* och *visst* får när de står i satsens mittfält och fundament.

Som exempel på studier som jämför språkanvändningen hos L1- och L2-talare på lexikal nivå kan nämnas bl.a. Hancock och Sanell (2009) och Pérez-Paredes (2010) som båda jämförde adverbbruket i L2- och L1-texter, Shimada (2014) som studerade användningen av diskursmarkörer i L2-engelska, och Ivaska (2014) som forskade i användningen av modalitetsuttryck i L2-finska. Användningen av konnektorer i L2-inlärarspråket har studerats av bl.a. Jiménez Catalán och Ojeda Alba (2010) som i sin studie *Connectors in EFL learners' essays and in course books* granskade användningen av konnektorer i L2-engelska. Informanterna i studien var spanska högstadiel elever från fyra olika skolor där fyra olika läroböcker på samma kunskapsnivå hade utnyttjats i undervisningen (Jiménez Catalán & Ojeda Alba, 2010: 87). Jiménez Catalán och Ojeda Alba kunde således jämföra

konnektorbruket i inlärttexter i förhållande till hur mycket exponering som informanterna hade fått för konnektorer i undervisningen. I resultaten kom det fram bl.a. att den lexikala variationen i användningen av konnektorer var låg i inlärttexter, och att de flesta informanterna använde högst två olika konnektorer i sin text. Den generella tendensen verkade överhuvudtaget vara att underanvända konnektorer, vilket enligt Jiménez Catalán och Ojeda Alba kan ha en negativ påverkan på hela den språkliga framställningen. (ibid.: 91.) Vad som Jiménez Catalán och Ojeda Alba (ibid.) dock fann ännu mer oroande var att exponeringen för konnektorer visade sig variera mycket i de olika läroböckerna, vilket betyder att inlärare i olika skolor inte överhuvudtaget hade haft jämlika förutsättningar för att lära sig använda konnektorer på korrekt och naturligt sätt.

Konnektorbruket i L2- och L1-engelska har även studerats av Granger och Tyson som i sin studie *Connector Usage in the English Essay Writing of Native and Non-Native EFL Speakers of English* (1996) utnyttjade CIA för att utreda om det verkar förekomma över- eller underanvändning av konnektorer i L1- och L2-engelska. Studiens resultat visade att det inte totalt sett verkade finnas någon tydlig tendens för överanvändning av konnektorer i L2-texter, men konnektorer som var förstärkande (*indeed, in fact, of course*), exemplifierande (*e.g., for instance, namely*) eller användes för att stödja argument med ytterligare uppgifter (*moreover*) verkligen överanvändes av andraspråksinlärare. Konnektorer som skapar kontrast (*however, though, yet*) eller som utvecklar argument (*therefore, thus*) däremot underanvändes i L2-materialet. (Granger & Tyson, 1996: 20.) En intressant iakttagelse i Granger och Tysons studie var användningen av uttrycket *on the contrary* som överanvändes av andraspråksinlärare som i studien var modersmålstalare av franska. Överanvändningen kunde enligt Granger och Tyson (ibid.: 22–23) förklaras med förstaspråkets påverkan: det franska uttrycket *au contraire* kan användas för att uttrycka såväl koncessiva som antitetiska relationer, och det går således att använda i delvis annorlunda kontexter än den engelska motsvarigheten.

En studie som delvis behandlar modala satsadverbial och som jag kommer att utnyttja när jag analyserar den här studiens resultat är Jenny Nilssons doktorsavhandling *Adverb i interaktion* (2005). I avhandlingen studerade Nilsson användningen av olika adverb i talad L1-svenska. Syftet var att finna och definiera de olika funktionerna som uttrycks med hjälp av adverb, samt beskriva vilka faktorer som påverkar de här funktionerna. Vidare ville Nilsson ta reda på vilka typer av sekvenser som adverb oftast förekommer i.

(Nilsson, 2005: 2–3.) Till studien valdes 48 adverb, varav en del var även satsadverbial (ibid.: 18–19).

Nilsson identifierade ett 30-tal olika huvudfunktioner för de 48 granskade adverb i sitt material. En del funktioner uppvisades av en mängd olika adverb, såsom kategorin *förmildrande* som realiserades av totalt tio olika adverb, medan några funktioner var unika för enskilda adverb, såsom kategorin *osannolikhet*, som bara realiserades av ordet *knappast*, och kategorin *generaliserande* som enbart uttrycktes med ordet *överhuvudtaget*. (Nilsson, 2005: 162–163.) Några adverb visade sig oavsett kontexten och positionen i satsen uttrycka bara en funktion, medan andra adverb kunde uppvisa flera olika funktioner, ibland till och med samtidigt (ibid.: 206). Trots att några av de granskade adverb kunde uttrycka flera olika funktioner, kom det inte fram några större överlappningar i funktioner. Endast ett par adverb delade exakt samma funktioner (t.ex. *ungefär* och *nästan* som båda kunde uttrycka såväl förmildrande som approximerande grad). Några enfunktionsadverb delade likaså densamma funktionen, t.ex. *antagligen* och *förmodligen* som båda uttryckte enbart trolighet i materialet. (ibid.: 207.) Nilsson (ibid.) sammanfattade tre huvudsakliga funktioner för adverb i materialet: att göra propositionen mindre exakt eller allmängiltig, att hålla med och visa känslomässigt engagemang i det man själv säger eller som reaktion till vad en annan samtalsdeltagare har sagt, och att uttrycka hurdan relation yttrandet har till samtalskontexten.

En del av funktionerna som Nilsson fann i materialet ansåg hon vara präglande för just talspråksmaterial, men de flesta funktionerna sammanfaller dock med tidigare beskrivningar av adverbens funktioner i skriftspråk (Nilsson, 2005: 206). Därför kommer jag att utnyttja Nilssons resultat även i min egen analys.

I kapitlen 2 och 3 har jag belyst den teoretiska bakgrunden för min undersökning genom att definiera de begrepp som är av central betydelse för studiens genomgång, samt genom att beskriva såväl den samhälleliga ramen som den akademiska ramen där studien ställer sig i. I följande kapitel vänds blicken mot själva studien.

4. MATERIAL OCH METOD

I det här kapitlet presenteras studiens material och metod. Som material i studien används korpusar, som diskuteras först allmänt i avsnitt 4.1, och sedan i avsnitt 4.2 ges en närmare översikt över de specifika korpusarna som används i den här undersökningen. Efter materialbeskrivningen redogör jag för studiens metodologi: i avsnitt 4.3 ges först en generell beskrivning över korpuslingvistik som forskningsmetod, och efter det presenteras CIA som formar den teoretiska grunden för min analys. Till slut beskriver jag närmare hur sökningar i materialet i praktiken har genomförts.

4.1 Korpusar

En korpus kan definieras som en samling av texter som nuförtiden ofta är i elektronisk form (Wray & Bloomer, 2012: 205). De flesta korpusarna innehåller skrivet språk eller transkriberat tal, men nuförtiden finns det även bl.a. teckenspråskorpusar (McEnery & Hardie, 2012: 3).

Gilquin, Granger och Meunier (2015: 1) talar om två primära fördelar som finns med användningen av korpusar i andraspråksforskningen. För det första är andraspråskorpusar ofta ganska stora och innehåller texter av ett stort antal inlärare. Därför är de ofta också representativa. För det andra, möjliggör den elektroniska formen att data kan analyseras med hjälp av ett antal olika verktyg. (ibid.) Att samla in ett forskningsmaterial som skulle vara tillräckligt omfattande och tillräckligt representativt, men också tillräckligt litet för att det är möjligt att analysera för hand på ett pålitligt och effektivt sätt skulle vara svårt. Att manuellt gå igenom och analysera material som innehåller miljoner ord skulle kräva mycket tid och resurser, och ändå kunde det vara omöjligt att göra det med tillräcklig noggrannhet på grund av mänskliga begränsningar. Även Lagerholm (2010: 48–49) konstaterar att det är tack vare korpusar och deras elektroniska format att ett flertal olika språkvetenskapliga ämnen, bl.a. grammatik, språkhistoria och textanalys, nuförtiden kan studeras på ett sätt som i praktiken inte var

möjligt tidigare. Det kan naturligtvis förekomma fel också med datorer, och t.ex. af Hällström-Reijonen (2014: 99–100) hävdar att åtminstone i språkforskning kan datorer aldrig ersätta det mänskliga ögat och den naturliga språkkänslan.

En enkel korpus kan bestå av ren text, men ju mer avancerad korpus, desto mer information kan ha inkluderats i den. En korpus kan vara *annoterad*, dvs. försedd med lingvistisk information, enligt bl.a. ordklass, lemgram⁵, syntaktisk struktur eller semantisk kategori (Wray & Bloomer, 2012: 205; van Rooy, 2015: 79). T.ex. ordklasstagning kan vara till stor nytta när det gäller ord som kan höra till flera ordklasser, som exempel nämner Wray och Bloomer (2012: 205) engelskans ord *record* som beroende av kontexten kan vara ett verb eller ett substantiv.

Det finns flera olika alternativ för hur man kan gå till väga med annoteringen. Det mest effektiva och tidsbesparande alternativet speciellt med större korpusar är att låta ett annoteringsprogram göra annoteringen automatiskt. Att annoteringen görs som manuellt arbete av en kunnig språkanalytiker kan däremot vara tids- och resurskrävande, men ger ofta mer tillförlitliga resultat. Ett tredje alternativ är att göra annoteringen delvis automatiskt och delvis manuellt. (van Rooy, 2015: 86–87.) McEnery och Hardie (2012: 30) poängterar att inte något av de här tre annoteringssätten är perfekt, och att det inte går att säga att den manuella annoteringen skulle vara det enda förfaringssättet som garanterar en tillförlitlig analys. Även en automatisk annotering av data kan i bästa fall producera till en hög grad felfri lingvistisk analys, medan en manuell annotering som utförs av en kunnig språkanalytiker inte nödvändigtvis leder till fullständigt felfria resultat på grund av mänskliga begränsningar. (ibid.) McEnery och Hardie (2012: 32) påpekar att korrektheten inte heller kan ses som det enda grundläggande kriteriet för en lyckad annotering, utan att det även är viktigt att annoteringen är konsekvent – ett kriterium som enligt McEnery och Hardie (ibid.) faktiskt mer sannolikt uppfylls hos automatiskt annoterade korpusar.

⁵ Ett *lemgram* står enligt Korps användarhandledning för ”ett ords eller flerordsuttrycks samtliga böjningsformer” (Språkbanken, u/å).

4.2 Material

Jag har valt att använda korpusar som material i undersökningen av två primära skäl. För det första utgör de korpusar som används ett relativt stort och representativt forskningsmaterial i förhållande till den tid och de resurser som är tillgängliga inom ramarna för den här undersökningen. För det andra är de också en praktiskt taget lämplig resurs: korpusarna är lätt tillgängliga på nätet och kan således användas i princip när som helst och var som helst.

L2-MATERIALET	L1-MATERIALET	
<i>”Inlärmaterialet”</i>	<i>”Läroboksmaterialet”</i>	<i>”Bloggmaterialet”</i>
SpIn, Tisus-texter, Sw1203- uppsatser	COCTAILL	Bloggmix 1998–2017, Bloggmix okänt datum

Tabell 2. Studiens material.

Som material i den här undersökningen används totalt 25 korpusar som jag har delat in i tre delmaterial. Det primära materialet, L2-materialet, som jag kallar för inlärmaterialet, består av tre inlärarkorpusar som presenteras i detalj i följande underkapitel. Inlärmaterialet jämförs med två typer av L1-material: ett läroboksmaterial och ett bloggmaterial, som också presenteras i korthet. Korpusarna är tillgängliga på nätet via Språkbanken⁶ som är en forskningsenhet vid Göteborgs universitet. Korpusarna bearbetas med korpusverktyget Korp (Borin, Forsberg & Roxendal, 2012). Härnäst presenterar jag de tre delmaterialen som används i undersökningen.

4.2.1 L2-materialet

Studios andraspråksmaterial består av tre korpusar: SpIn, Tisus-texter och Sw1203-uppsatser.⁷ Tillsammans utgör de ett material på 156 988 tokens.⁸ De här tre korpusarna

⁶ <https://spraakbanken.gu.se/korp>

⁷ Det finns flera tillgängliga versioner av inlärarkorpusarna i Korp. I den här undersökningen används de versioner som finns i kategorin *Skyddade korpusar*: SpIn, Sw1203-uppsatser version 2 och Tisus-texter v2, men för enkelhetens skull kallas de två sistnämnda i den här avhandlingen för Sw1203 och Tisus.

⁸ Ett *token* står för individuella förekomster av enskilda ordtyper, inkluderande även t.ex. skiljetecken (Sigurd & Håkansson, 2007: 188; Språkbanken, u/å).

är skyddade, vilket betyder att de inte är fritt tillgängliga för allmänheten. Jag har fått tillgång till korpusarna för att genomföra den här undersökningen.

Korpus	Antal texter	Antal tokens	Antal meningar
SpIn	244 essäer / 160 studenter	44 996	4 124
Sw1203	141 essäer / 54 studenter	51 956	3 254
Tisus	105 essäer / 105 studenter	60 036	3 708
Totalt	490 essäer / 319 studenter	156 988	11 086

Tabell 3. Numeriska data om inlärmaterialet (beräknat i Korp).

SpIn-korpusen består av 244 essäer skrivna av 160 studenter på en språkkurs på Centrum för Språkintröduktion under åren 2013–2015. Programmet är avsett för unga, 16–20 år gamla invandrare som en förberedande kurs för fortsatta studier i offentliga skolor i Sverige. (Volodina m.fl. 2016: 207.) De kunskapsnivåer som är representerade i korpusen går från nivå A1 till nivå B2.

Sw1203-korpusen innehåller 141 essäer som har skrivits av 54 studenter på kursen *Swedish as a foreign language – Qualifying course in Swedish* som arrangerades i 2012–2013 vid Göteborgs universitet. Kursen är avsedd för utländska studenter, och den syftar på att ge kvalifikation för universitetsstudier i Sverige. Varje student har skrivit tre essäer, en i början av kursen, en i mitten och en i kursens slutprov. (Volodina m.fl., 2016: 207.) Kunskapsnivåerna går från nivå B1 till nivå C2.

Texterna i Tisus-korpusen är insamlade i det så kallade Tisus-provet (*Test in Swedish for University Students*). Genom att delta i provet kan studenten bli kvalificerad för att söka till universitetsstudier i högskolor i Sverige, och själva provet består av såväl skrivna uppgifter som ett muntligt prov och en läsförståelsetest. De 105 essäer som har samlats in i Tisus-korpusen är skrivna av 105 studenter år 2006–2007, och på CEFR-skalan representerar de nivåerna B2 och C1. (Volodina m.fl., 2016: 207.)

	A1	A2	B1	B2	C1	C2	Okänd
SpIn	55	137	44	2	0	0	6
Sw1203	0	0	40	71	23	7	0
Tisus	0	0	0	32	73	0	0
Totalt	55	137	84	105	96	7	6

Tabell 4. Antalet texter på olika CEFR-nivåer i inlärmaterialet (beräknat i Korp).

	A1	A2	B1	B2	C1	C2	Okänd
Tokens	5172	26 200	23 921	45 182	51 549	3 258	1 706

Tabell 5. Antalet tokens på olika CEFR-nivåer i inlärmaterialet (beräknat i Korp).

De tre inlärkorporusarna är automatiskt annoterade och försedda med metadata om informanterna och texterna (t.ex. kön, ålder, modersmål, uppsatsämne, datum). Varje informant och varje text har ett eget identifikationsnummer som gör det möjligt att skilja olika informanter och texter åt. Texterna i Sw1203 och SpIn har bedömts och placerats på CEFR-skalan av kvalificerade språklärare. (Volodina m.fl. 2016: 208–209.) I Tisus-korpusen är CEFR-nivåerna snarare riktgivande – enligt korpusmetadata har texterna betygsatts i skalan 1–5, och texterna som fått betyget 3 eller högre har markerats representera nivån C1, medan texterna som fått betyget 1 eller 2 har markerats representera nivån B2.

Användningen av inlärkorporusar som forskningsmaterial förutsätter att man tar hänsyn till den etiska aspekten som finns med det här materialvalet – inlärkorporusar innehåller ju texter som inte är på samma sätt avsedda för allmänheten som t.ex. lärobokstexter och bloggtexter. Därför är det för det första viktigt att konstatera att informanterna har gett sitt samtycke till att texterna får användas i forskning. När texterna har sparats i Korp har all känslig information, t.ex. namn och gatuadress, pseudonymiserats, vilket säkerställer att informanterna inte kan igenkännas av texterna. (Volodina m.fl., 2016: 208.) I min undersökning behandlas informanterna dessutom främst på gruppnivå och ingen onödig information om enskilda informanter avslöjas i undersökningen – om det t.ex. är nödvändigt att skilja olika individer åt, används det personliga informantnumret. Personuppgifter som är tillgängliga på Korp, såsom informantens ålder, kön eller modersmål, är inte relevanta med tanke på den här undersökningens syfte, och den informationen inkluderas således inte i analysen. Eftersom inlärkorporusarna är som sagt skyddade och inte i sig tillgängliga för allmänheten, har jag även kontrollerat av Språkbanken att det är tillåtet att citera materialet i form av exempelmeningar i den här avhandlingen.

4.2.2 L1-materialet

Studiens L2-material jämförs med två typer av förstaspråksmaterial. Först jämförs inlärartexterna med lärobokstexter för att se vilka modala satsadverbial som andraspråksinlärare överhuvudtaget exponeras för i undervisningen. Studiens läroboksmaterial formas av Språkbankens COCTAILL-korpus som innehåller texter från 12 läroböcker i svenska som andraspråk (Volodina m.fl., 2014: 128). Läroböckerna har valts så att de motsvarar kunskapsnivåerna A1, A2, B1, B2 och C1. Nivån C2 är inte representerad i korpusen eftersom inga läroböcker på nivå C2 var tillgängliga när korpusen samlades. En del av böckerna täcker fler än en nivå, och totalt representeras nivån C1 av två olika böcker, och alla andra nivåer representeras av fyra olika böcker var och en. (Volodina m.fl., 2014: 131–132.) COCTAILL är en fritt tillgänglig korpus på Språkbanken, och den är automatiskt annoterad.

CEFR-nivå	Antalet böcker	Antalet tokens
A1	4	34 464
A2	4	80 677
B1	4	147 363
B2	4	161 740
C1	2	170 043
Totalt	18 (12 enskilda böcker)	710 251

Tabell 6. Numeriska data om läroboksmaterialet (Volodina m.fl., 2014: 132; Korp).

Den andra delen av L1-materialet, som jag kallar för bloggmaterial, består av sammanlagt 21 enskilda Bloggmix-korpusar. De här korpusarna används som jämförelsematerial för att se hur användningen av modala satsadverbial hos andraspråksinlärare förhåller sig till det faktiska språkbruket hos modersmålstalare. Bloggmix-korpusarna innehåller texter från svenska bloggar från åren 1998–2017⁹, och de är fritt tillgängliga på Språkbanken. Såsom inlärarkorpusarna och COCTAILL-korpusen, är Bloggmix-korpusarna också automatiskt annoterade.

En blogg är enligt Svensk Ordbok (2021) en ”dagbok eller forum på nätet som uppdateras kontinuerligt och är öppen för kommentarer”. Enligt Kaye (2010: 196) var bloggar till en början ofta en slags dagböcker där den avsedda läsarkretsen bestod av skribentens vänner

⁹ En av korpusarna innehåller texter med ett okänt datum.

och familjemedlemmar, medan nuförtiden kan bloggar också ägnas åt specifika teman, såsom t.ex. matlagning, resande eller inredning, och de är ofta avsedda för större publik. Bloggar skrivs såväl av privatpersoner ("bloggare") som representanter för företag och organisationer, och de kan förutom text innehålla t.ex. foton och andra grafiker. (Kaye, 2010: 194, 196.) Speciellt i personliga bloggar är språket enligt Lövheim (2011: 8) ofta rätt informellt och personligt. Vad som skiljer bloggar från mer traditionella mediakällor är särskilt det interaktiva formatet. Trots att själva bloggtexten skrivs av bloggaren, kan även läsare delta i diskussionen genom att kommentera texten eller genom att svara på andra läsaes kommentarer. (Kaye, 2010: 194.) På detta sätt kan bloggtexterna beskrivas som en mer dialogisk än monologisk texttyp.

Användningen av bloggtexter som jämförelsematerial för L2-inlärttexter är inte ideal men inte heller omotiverad. Det är självklart att i det bästa fallet skulle inlärtmaterialet jämföras med ett exakt motsvarande data från L1-talare, men sådant material var tyvärr inte tillgängligt när den här undersökningen genomfördes. Efter ett noggrant övervägande valdes bloggtexter som jämförelsematerial av tre primära skäl. För det första kan bloggtexter ses som en varierande genre. De 21 Bloggmix-korpusarna kan antas innehålla rätt olika typer av skrifter som behandlar olika typer av ämnen ur olika typer av perspektiv. De präglas alltså inte alltför mycket av för texttypen utmärkande språkliga särdrag, något som kunde vara fallet med någon mer homogen texttyp. Dessutom anser jag att bloggtexter inte representerar på samma sätt professionellt språk som t.ex. myndighetstexter eller skönlitterära texter. Bloggtexter skrivs ofta av helt vanliga människor, och språkbruket är således naturligt och inte lika formellt, vårdat eller genomtänkt som det kan vara i mer professionellt laddade texttyper. För det andra, är syftet med bloggar enligt Lagerholm att "offentliggöra sina tankar och åsikter om det som engagerar skribenten" (2008: 353). Därför är det antagligt att åtminstone en del av texterna i bloggmaterialen är ställningstagande, till och med argumenterande, och de kan således vara överraskande likartade med studiens L2-material, som för det mesta består av argumenterande essäer. Den tredje motiveringen för materialvalet är att de totalt 21 Bloggmix-korpusarna tillsammans utgör ett material på 615 miljoner tokens. De kan förväntas således ge en rätt täckande syn på det verkliga språkbruket hos modersmålstalare.

4.3 Metod

Korpuslingvistik är en språkvetenskaplig forskningsmetod som använder sig av korpusar. Metoden skiljer sig från många andra språkvetenskapliga metoder i det att det inte är en enda metod som fokuserar på en viss aspekt av språk, utan att det faktiskt är fråga om ett flertal metoder. Detta för sin del gör det möjligt att studera flera olika typer av forskningsfrågor ur ett korpuslingvistiskt perspektiv. (McEnery & Hardie, 2012: 1.) Eftersom korpuslingvistiken inte kan ses som en enda metod har några forskare valt att definiera den snarare som t.ex. verktyg, teori, forskningsfält eller paradigm (se t.ex. Taylor, 2008), men i den här avhandlingen behandlas korpuslingvistiken som forskningsmetod.

McEnery och Hardie (2012: 1–2) definierar korpuslingvistiken utifrån fyra generaliseringar som kan anses känneteckna de flesta korpuslingvistiska studier. För det första använder korpuslingvistiken sig av (numera elektroniska) textsamlingar som ofta är rätt stora. På grund av den stora mängden data är det inte lätt att bearbeta material på ett pålitligt och effektivt sätt för hand. Den andra generaliseringen lyder således att man i korpuslingvistiska undersökningar ofta utnyttjar olika elektroniska verktyg för att bearbeta material. (McEnery & Hardie, 2012: 2.) De två följande generaliseringar framhäver ett noggrant val av korpusar. För att kunna utföra en korpuslingvistisk undersökning ska man alltid tänka på ändamålet: det är ytterst viktigt att den korpusen som används i en undersökning lämpar sig för forskningsfrågorna. (ibid.) Som den fjärde och sista generaliseringen lyfter McEnery och Hardie (ibid.) fram det faktumet att det är viktigt att vara medveten om hur texterna i en korpus inte nödvändigtvis representerar helt homogena texttyper. Till exempel kan en lärobokskorpus bestå av texter från olika delar av en lärobok (t.ex. dialoger, grammatikbeskrivningar, olika typer av övningar) som kan utmärkas av rätt olika språkliga särdrag, vilket kan leda till snedvridande resultat.

4.3.1 Korpustypologi

McEnery och Hardie (2012: 3) presenterar en grov uppdelning av kriterier som kan användas för att skapa en typologi för korpuslingvistisk forskning. Typologin består av

sex delar. Härnäst presenterar jag de sex delområdena och diskuterar hur korpusarna i min undersökning förhåller sig till dem.

1. *Kommunikationssätt*. Med kommunikationssätt (*mode of communication*) avses den språkliga formen som finns dokumenterad i korpusen. I det enklaste fallet står kommunikationssättet alltså för distinktionen mellan talat och skrivet språk, men nuförtiden är även korpusar av teckenspråk, gester eller andra paralingvistiska drag tillgängliga. (McEnery & Hardie, 2012: 3.) I den här undersökningen representerar alla tre delmaterialen skrivet språk.

2. *Korpusbaserade vs korpusdrivna studier*. Den följande distinktionen kan göras mellan korpusbaserade och korpusdrivna studier (*corpus-based vs corpus-driven studies*). I en korpusbaserad studie används korpusar för att undersöka en teori, hypotes eller forskningsfråga, medan i en korpusdriven studie används korpusen som utgångspunkt för formuleringen av hypotes eller metod. (McEnery & Hardie, 2012: 6.) Den här undersökningen är korpusbaserad, eftersom korpusar används för att undersöka användningen av modala satsadverbial. En korpusbaserad undersökning möjliggör en omfattande och tillförlitlig analys av det verkliga språkbruket – något som t.ex. Biber m.fl. (1998: 4) hävdar inte ens är möjligt på något annat sätt.

3. *Materialinsamling*. Vidare kan korpuslingvistiska undersökningar kategoriseras enligt det förfaringssättet som används för materialinsamlingen (*data collection regime*). McEnery och Hardie (2012: 6) talar om *monitor corpora*, eller monitorkorpusar, som är korpusar där nytt material läggs till ständigt. Monitorkorpusar kan innehålla rätt olika typer av material, och den ständiga tillväxten kan leda till en situation där några materialtyper är betydligt mycket mera eller mindre representerade i korpusen än andra material. De fördelar som finns med monitorkorpusar överskrider det här problemet: även de möjliga skevheter i materialet blir korrigerade av sig själv då korpusen växer. (McEnery & Hardie, 2012: 6–7.) Den andra typen av korpusar kan kallas för *balanserade korpusar* (*balanced* eller *sample corpora*). Materialet i balanserade korpusar representerar en särskilt typ av språk i en särskild tidpunkt. Syftet är således att materialet skulle vara så representativt och balanserat som möjligt i en given samplingsram (*sampling frame*). Ett exempel på samplingsram kunde vara t.ex. servicesamtal i affärer på 1990-talet i Storbritannien – i det här fallet kunde bara ett noga samlat och begränsat urval av material användas i korpusen. Fördelen med balanserade korpusar är att de kan

ge en mer preciserad syn på språk, men för att uppnå representativa och välbalanserade resultat är det nödvändigt att materialet väljs ut med omsorg. (McEnery & Hardie, 2012: 8–9.) Korpusarna i min undersökning är balanserade, eftersom de innehåller texter av en särskild typ och från en specifik tidpunkt. De är inte heller växande, men de kan uppdateras.¹⁰

4. *Annoterade vs oannoterade korpusar.* Den fjärde distinktionen gör McEnery och Hardie (2012: 13) mellan annoterade och oannoterade korpusar (*annotated vs unannotated corpora*). En viktig fördel med annoterade korpusar är att de ger forskaren en snabb tillgång till grammatisk information, såsom t.ex. grammatiska kategorier, men McEnery och Hardie konstaterar att annoterade korpusar har också sina motståndare:

[--] arguments against annotation are largely predicated upon the **purity of the corpus texts** themselves, with the **analyses being viewed as a form of impurity**. This is because they impose an analysis on the users of the data, but also because **the annotations themselves may be inaccurate or inconsistent**. (McEnery & Hardie, 2012: 14.) (Min emfas)

Valet mellan annoterade och oannoterade korpusar är alltså en vattendelare mellan olika korpuslingvistiska skolor. Några lingvister jobbar hellre med oannoterade korpusar för att undvika de möjligen felaktiga eller inkonsekventa annoteringarna och för att kunna själv skapa den lingvistiska analysen från början. McEnery och Hardie (2012: 14) påpekar att medan användningen av oannoterade korpusar verkligen lämnar utrymme för forskarens egen analys, är risken då att de möjliga fel i resultat inte kan spåras till materialet på samma sätt som de kan i en annoterad korpus.

I den här undersökningen används automatiskt annoterade korpusar, och det är möjligt att räkna upp många fördelar med det valet. Det är inte bara effektivt med hänsyn till de tillgängliga resurserna, utan säkerställer också att resultaten blir enhetliga. Tack vare den genomskinliga analysprocessen är resultaten dessutom lätta att reproducera vilket ökar dess reliabilitet. Trots de här fördelarna är jag medveten om de möjliga fel och bristfälligheter som kan orsakas av den automatiska annoteringen. Därför kommer jag enligt den tredje forskningsfrågan i samband med resultaten att diskutera och överväga

¹⁰ T.ex. inlärartexterna har omgraderats och kunskapsnivåerna i inlärarkorpusarna har således blivit uppdaterade.

materialvalet för att se om det är de obestridliga fördelarna eller de möjliga nackdelarna som till slut väger tyngst.

5. *Urval av data*. Den femte observationen som McEnery och Hardie (2012: 14) tar upp handlar om hur data väljs ut (*total accountability vs data selection*). En forskare som använder korpusar för att testa en specifik teori kan omedvetet välja ut en sådan del av data som stödjer hypotesen, vilket naturligtvis inte ger reliabla resultat. Ett sätt att undvika sådana förvrängningar i resultaten är att använda korpusar i sin helhet. Om det inte är möjligt att använda korpusar i sin helhet och någon slags selektion av data måste göras, ska den göras så att man inte avsiktligt väljer ut sådana exempel som stödjer hypotesen eller lämnar ut sådana som inte gör det. (McEnery & Hardie, 2012: 14–15.) I den här undersökningen används totalt 25 korpusar, alla i sina helheter, och alla resultat tas i beaktande.

6. *Flerspråkiga vs enspråkiga korpusar*. Den sjätte och sista distinktionen som presenteras av McEnery och Hardie (2012: 18) gäller användningen av flerspråkiga korpusar mot enspråkiga korpusar (*monolingual vs multilingual corpora*). En korpus kan sägas vara flerspråkig om det finns minst två språk som systematiskt förekommer i korpusen (t.ex. parallellkorpusar). En enspråkig korpus innehåller i princip bara ett språk, men i praktiken kan dock en enspråkig korpus innehålla enstaka ord eller fraser på andra språk. Det väsentliga är vad det primära syftet med korpusen är, dvs. vilket språk den avses att representera. (McEnery & Hardie, 2012: 18–19.) Korpusarna i min undersökning är enspråkiga, för att de först och främst beskriver svenska språket.

4.3.2 Contrastive Interlanguage Analysis

För att se hur användningen av modala satsadverbial skiljer sig mellan L1- och L2-svenska, och vidare mellan andraspråksinlärare på olika kunskapsnivåer, utnyttjar jag Sylviane Grangers *contrastive interlanguage analysis* (CIA).

Det finns olika sätt att tillämpa CIA i andraspråkforskning (se Granger, 2015), men ett vanligt sätt att utnyttja CIA är att utgå från två typer av jämförelser. Å ena sidan jämförs språkanvändningen mellan modersmålstalare (*NS: native speaker*) och icke-modersmålstalare (*NNS: non-native speaker*), och å andra sidan jämförs språkanvändningen mellan olika icke-modersmålstalargrupper, t.ex. mellan människor

med olika modersmål (se bild 2). Jämförelsen mellan modersmålstalare och icke-modersmålstalare syftar på att lyfta fram sådana drag i inlärarespråket som avviker från den målspråksenliga användningen av språket, t.ex. över- eller underanvändning av ord, fraser eller strukturer. (Granger, 2002: 12.) Att jämföra språkanvändning mellan flera icke-modersmålstalare kan däremot berätta mer om inlärarespråket och visa vilka av de avvikande dragen som verkar universellt höra till inlärarespråket och vilka som är kännetecknande för t.ex. en viss nationalitetsgrupp och kan möjligtvis berätta om överföring från ett visst förstaspråk (*transfer*) (Granger, 2002: 13). I den här undersökningen jämförs språkanvändningen mellan andraspråksinlärare och förstaspråkstalare (NS vs NNS) och mellan andraspråksinlärare på olika kunskapsnivåer (NNS vs NNS).

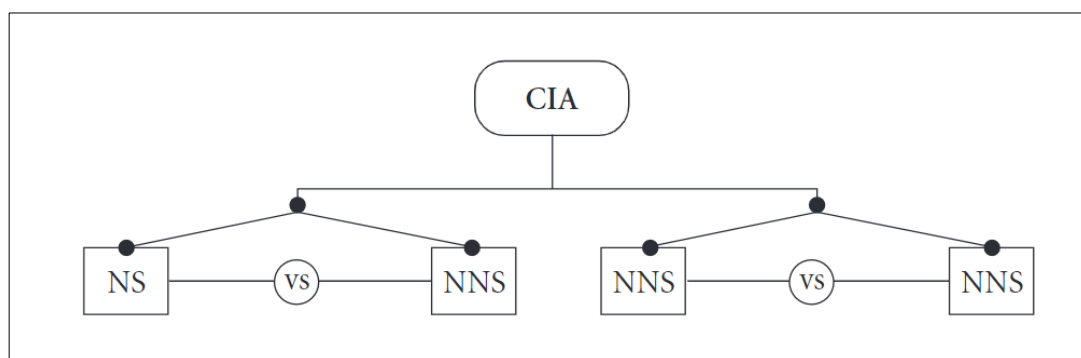


Bild 2. *Contrastive Interlanguage Analysis* enligt Granger (2002: 12, inklusive bilden).

4.3.3 Sökningarna i Korp

Sökningarna i Korp görs med hjälp av korpusverktygets utökade sökning.¹¹ För att hitta alla modala satsadverbial i materialen utnyttjar jag annoteringen enligt den syntaktiska relationen (dependensrelation) och söker efter alla ord och uttryck vars dependensrelation har markerats vara satsadverbial. Trots att bestämningen *modala* saknas i namnet, visar den här sökningen ord som annoteringsprogrammet har markerat med tagget MA (*Attitude adverbial*), en kategori som består av modala satsadverbial.¹² Bara de

¹¹ Sedan den här undersökningen inleddes har det skett förändringar i Korp, och de sökningar som jag har gjort för att producera resultat motsvarar således inte nödvändigtvis sökningar som skulle vara möjliga i den nuvarande versionen av Korp.

¹² Annoteringen i Korp baserar sig på Telemans (1974). Inom ramarna för den här avhandlingen är det inte möjligt att gå närmare in på de olika kategorierna, men ord som har taggats med MA motsvarar kategorin som Telemans (1974: 89) kallar för *talorattitydsadverbial*, i praktiken alltså modala satsadverbial.

satsadverbial som används av fler än en informant inkluderas i analysen för att säkerställa att resultaten är representativa på gruppnivå.

För att tolka resultaten används korpusverktygets statistik-flik som visar förekomsterna av olika satsadverbial samt deras absoluta och relativa frekvenser. Det är möjligt att sammanställa resultaten enligt olika kriterier, vilket utnyttjas i undersökningen. En lista över alla sökningar som genomförts för att producera resultaten finns samlade in i bilaga 1 i slutet av avhandlingen, så att resultaten är vid behov lätta att reproducera.

De absoluta antalen av enskilda satsadverbial kan inte användas för att jämföra användningen av modala satsadverbial på olika kunskapsnivåer eftersom alla kunskapsnivåer inte är lika representerade i materialet (se tabellerna 4 och 5). Därför är det förväntat att de absoluta antalen satsadverbial är större t.ex. på nivå B1 än på nivå C2, vilket dock inte betyder att satsadverbial nödvändigtvis skulle användas mer på nivå B1. För att kunna jämföra olika nivåer sinsemellan ska man därför räkna de olika satsadverbialens relativa frekvenser per nivå. Den relativa frekvensen normaliserar antalet träffar till en miljon ord och berättar således hur många gånger ordet kan förväntas förekomma i en löpande text på en miljon ord.¹³ De relativa frekvenserna (Y) av olika modala satsadverbial per nivå räknas med följande formel där x är det totala antalet träffar av ett satsadverbial och z är det totala antalet tokens i texter på respektive nivå:

$$Y = (x/z) * 10^6$$

¹³ McEnery och Hardie (2012: 50) påpekar att inom korpuslingvistik är det vanligast att normalisera till ett tusen eller en miljon ord. I den här avhandlingen har jag valt att normalisera till en miljon ord på grund av Bloggmix-materialets storlek.

5. RESULTAT OCH ANALYS

I det här kapitlet går jag igenom undersökningens resultat. Först behandlas resultaten ur kvantitativt perspektiv genom att titta på vilka modala satsadverbial som förekommer i studiens tre delmaterial, först i inlärttexter och lärobokstexter (5.1 och 5.2), och sedan i bloggtexter (5.3). Efter den kvantitativa jämförelsen gör jag även en kort kvalitativ analys av resultaten i avsnitt 5.4. Till slut, i avsnitt 5.5, granskar jag med hjälp av en kontrollkorpus hur den automatiska annoteringen har påverkat studiens resultat.

5.1 Kvantitativ analys av resultaten

En sökning efter belägg som taggats som satsadverbial¹⁴ i inlärmaterialet ger totalt 555 träffar. Den relativa frekvensen för satsadverbial i hela inlärmaterialet är således 3535,3. En närmare titt på beläggen visar att en stor del av de ord och uttryck som har klassificerats som satsadverbial i materialet används mindre än 5 gånger, vilket inte ger en omfattande syn på användningen av de här satsadverbialen på gruppnivå. Därför är det befogat att fokusera på den begränsade gruppen av satsadverbial som är de mest frekventa i materialet.

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	141	898,2
Ju	48	305,8
Verkligen	40	254,8
Egentligen	28	178,4
Naturligtvis	26	165,6

Tabell 7. De fem mest frekventa modala satsadverbialen i hela inlärmaterialet.¹⁵

¹⁴ Hädanefter i avhandlingen talar jag om *satsadverbial* i betydelsen ”ord som annoterats med tagget MA” – dvs. modala satsadverbial. För de andra typer av satsadverbial hänvisar jag till vid behov med en särskild bestämning (t.ex. ’negerande satsadverbial’).

¹⁵ Sökning 1 (bilaga 1)

De fem mest frekventa modala satsadverbialen i hela inläromaterialet presenteras i tabell 7. Det framgår av tabellen att det allra mest frekventa modala satsadverbial i materialet är *kanske* som förekommer totalt 141 gånger. Det betyder att av de 555 träffarna av satsadverbial i materialet utgör *kanske* totalt ca 25 % av förekomsterna. På andra plats följer *ju* med 48 träffar, dvs. ca 9 % av alla belägg. Efter *ju* följer *verkligen* med 40 träffar (ca 7 %), *egentligen* med 28 träffar (ca 5 %) och *naturligtvis* med 26 träffar (ca 5 %). De fem mest frekventa modala satsadverbialen utgör således över hälften (ca 51 %) av alla satsadverbialen i inläromaterialet.

Ett delsyfte i avhandlingen är att evaluera hur användningen av automatiskt annoterade korpusar lämpar sig för den här typen av undersökning. Ett sätt att göra det är att jämföra resultat som produceras automatiskt av korpusverktyget med resultat som fås genom att manuellt söka enskilda ord i materialet. Efter att ha rätt ut vilka som är de fem mest frekventa satsadverbialen i inläromaterialet enligt den automatiska annoteringen, sökte jag efter de fem orden var för sig för att se hur många träffar av orden i verkligheten finns i materialet. Resultaten har samlats in i tabell 8.

Satsadverbial	<i>Af</i> (taggat som satsadverbial)	<i>Af</i> (alla träffar)	Δ
Kanske	141	178	-37
Ju¹⁶	48	50	-2
Verkligen	40	60	-20
Egentligen	28	65	-37
Naturligtvis	26	30	-4

Tabell 8. De fem mest frekventa modala satsadverbialen i inläromaterialet, taggat som satsadverbial vs alla träffar.¹⁷

Det framgår av tabellen att antalen träffar av de fem mest frekventa modala satsadverbialen i materialet är större när sökningen inte beaktar dependensrelationen. Den största skillnaden finns med ordet *egentligen*: av de totala 65 träffarna har över hälften (ca 57 %) taggats som någonting annat än satsadverbial. *Ju* har däremot nästan alltid taggats som satsadverbial (96 % av träffarna) och så är också ordet *naturligtvis* (ca 87 %). Att de absoluta antalen inte motsvarar varandra kan bero på att orden förekommer också som andra satsdelar i materialet men det kan också tyda på att en del av förekomsterna är

¹⁶ Jag har räknat bort sådana fall där *ju* förekommer i fraserna *ju-desto* eller *ju-ju*.

¹⁷ Sökning 2 (bilaga 1)

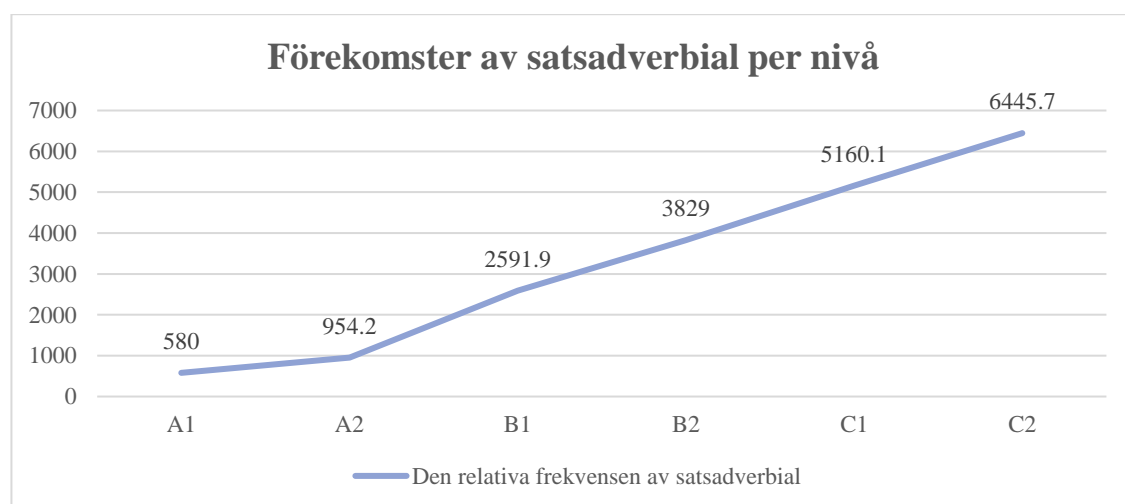
felaktigt taggade. De här resultaten ger i alla fall anledning för en närmare granskning av annoteringens kvalitet, något som jag kommer att göra i ett eget avsnitt (se avsnitt 5.5).

5.2 Användningen av modala satsadverbial på olika kunskapsnivåer

Efter att ha granskat hur modala satsadverbial överhuvudtaget används i inlärttexter, är det dags att vända blicken mot hur dem används på olika kunskapsnivåer. Numeriska data om användningen av modala satsadverbial på olika kunskapsnivåer presenteras i tabell 9. En lista över samtliga ord som har taggats som satsadverbial på de olika kunskapsnivåer finns i bilaga 2 i slutet av avhandlingen.

Nivå	Absolut frekvens	Relativ frekvens
A1	3	580,0
A2	25	954,2
B1	62	2 591,9
B2	173	3 829,0
C1	266	5 160,1
C2	21	6 445,7

Tabell 9. Frekvenserna av tokens som annoterats som satsadverbial i inlärmaterialet på olika kunskapsnivåer.¹⁸



Figur 1. Den relativa frekvensen av tokens som annoterats som satsadverbial på olika kunskapsnivåer i inlärmaterialet.

¹⁸ Sökning 3 (bilaga 1)

Såsom det framgår av tabell 9 och figur 1, ökar användningen av modala satsadverbial ju högre upp studenten går på CEFR-skalan. Figur 1 visar att användningen växer relativt stabilt och inga branta stigningar sker mellan nivåerna. Det kan tolkas som ett tecken på att användningen av modala satsadverbial ökar jämnt ju mer avancerad nivå studenten når. Att användningen av modala satsadverbial ökar medan språkkunskaper också i övrigt förbättras är väntat, men det är intressant hur linjär tillväxten är.

I de siffrorna som berättar om användningen av modala satsadverbial i tabell 9 och figur 1 har korpusverktyget räknat med alla belägg som har taggats som satsadverbial i materialet. I siffrorna ingår alltså även ord som inte är just modala satsadverbial, eller inte ens satsadverbial. Dessutom kan några belägg saknas i resultaten eftersom de inte har taggats som satsadverbial. Frekvenserna är alltså ungefärliga, men eftersom det var rätt förväntat att modala satsadverbial skulle användas mer när språkkunskaperna också i övrigt förbättras, kan den stigande figuren ändå tolkas som riktgivande. I följande avsnitt går jag igenom resultaten per nivå och granskar närmare vilka ord som har taggats som satsadverbial på enskilda nivåer.

5.2.1 Modala satsadverbial på nybörjarnivå

I den delen av inläromaterialet som representerar den så kallade nybörjarnivån, dvs. nivåerna A1 och A2, förekommer väldigt få satsadverbial. En sökning efter ord som har taggats som satsadverbial ger 3 träffar på nivå A1 och 25 träffar på nivå A2.¹⁹ En titt på vilka ord som i själva verket formar de här siffrorna visar att det faktiskt inte används några modala satsadverbial på nivå A1, utan de ord som har taggats som satsadverbial är felannoterade. På nivå A2 är det bara *kanske* som systematiskt används i inlärotexterna: av de 25 belägg som har taggats som satsadverbial på nivå A2 är 11 förekomster av *kanske*. De 14 andra orden som Korp har kategoriserat som satsadverbial innehåller två andra modala satsadverbial, nämligen *tyvärr* och *förstås*, som båda förekommer bara en gång i texterna. Att det enda modala satsadverbial som överhuvudtaget används av fler än en skribent på nivå A2 är *kanske* visar att den lexikala variationen i materialet är låg, åtminstone när det gäller modala satsadverbial. Att även *kanske* har en rätt låg relativ frekvens på nivå A2 i materialet (rf=419,8), jämfört med t.ex. det negerande

¹⁹ Sökning 4 (bilaga 1)

satsadverbialen *inte*²⁰ som redan på nivå A2 har en relativ frekvens på 10 305,3, kan för sin del möjligen tyda på att *inte* ens *kanske* har i större utsträckning etablerat sin plats i ordförrådet hos andraspråksinlärare på nybörjarnivå, trots att det ändå är det enda modala satsadverbialen som någorlunda systematiskt används i inlärtexterna på nivå A2.

Den låga frekvensen av modala satsadverbial på de två lägsta nivåerna är inte överraskande eftersom språkinlärningsprocessen är i sin början, och speciellt det produktiva ordförrådet kan förväntas vara knappt och innehålla bara de mest grundläggande ord, såsom t.ex. ordet *inte* som i inlärtexterna redan på nivå A1 används totalt 44 gånger och har således en relativ frekvens på 8507,3.²¹ Även CEFR-skalan beskriver att ordförrådet på nivå A1 består av enstaka ord och fraser som ”rör vissa konkreta situationer” (Skolverket, 2009: 109), vilket naturligtvis i sig utesluter de modala satsadverbial som har en mer abstrakt betydelse.

Att modala satsadverbial inte ännu på nybörjarnivån i större utsträckning tycks ha blivit etablerade i det produktiva ordförrådet, dvs. att informanten inte själv använder orden i skriftlig produktion, betyder dock inte att informanterna inte skulle kunna känna till och förstå orden när de används i skrift eller tal. En sökning efter satsadverbial på nivåerna A1 och A2 i COCTAILL-korpusen visar att andraspråksinlärare introduceras till några modala satsadverbial i läroböckerna redan på nivåerna A1 och A2, och det är därför troligt att inlärnarna senast på nivå A2 faktiskt har ett större antal modala satsadverbial i sitt receptiva ordförråd än vad de själva använder i skrift. De mest frekventa modala satsadverbialen i läroböckerna på nivåerna A1 och A2 presenteras i tabellerna nedan.

Satsadverbial	Absolut frekvens	Relativ frekvens
Ju	7	203,1
Kanske	5	145,1
Gärna	3	87,1
Väl	2	58,0
Naturligtvis/förstås/faktiskt	1	29,0

Tabell 10. De fem mest frekventa modala satsadverbialen på nivå A1 i COCTAILL.²²

²⁰ Sökning 5 (bilaga 1)

²¹ Sökning 5 (bilaga 1)

²² Sökning 6 (bilaga 1)

Tabell 10 visar att modala satsadverbial inte är särskilt ofta förekommande i läromaterialen på nivå A1. Den högsta frekvensen är med ordet *ju*, men även dess relativa frekvens är mycket låg jämfört med t.ex. ordet *inte* som har en relativ frekvens på 6 354,5 i lärobokstexterna på nivå A1.²³ Överraskande är hur det modala satsadverbial som inlärares mest exponeras för i läromaterialen på nivå A1 är just *ju* och *inte kanske*, eftersom *kanske* ändå är det enda modala satsadverbial som används i inlärares texterna redan på nivå A2. Som det framgår av de absoluta frekvenserna, är skillnaden mellan frekvenserna av *ju* och *kanske* dock i själva verket inte stor.

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	51	632,2
Ju	50	619,8
Gärna	35	433,8
Verkligen	11	136,3
Väl/förstås	10	124,0

Tabell 11. De fem mest frekventa modala satsadverbialen på nivå A2 i COCTAILL.²⁴

På nivå A2 får studenten mer exponering för modala satsadverbial i läromaterialen. *Kanske* är det mest vanligt förekommande modala satsadverbial i materialet, och *ju* följer på andra plats. Vartdera ord används ca 50 gånger i texterna och deras relativa frekvenser är ca 632 och 619 respektive, vilket är som förväntat mycket mera än de motsvarande talen i läromaterialen på nivå A1. De följs av ordet *gärna* med en relativ frekvens på ca 433. Orden *verkligen* (rf=136,3) samt *väl* och *förstås* (rf=124,0) står på fjärde och femte plats, men deras relativa frekvenser är rätt mycket lägre och således är de mer marginella. Det är ändå tydligt att modala satsadverbial har börjat synas mera i lärobokstexter, och att det kan förväntas att de så småningom också börjar dyka upp i inlärares egen textproduktion.

²³ Sökning 7 (bilaga 1)

²⁴ Sökning 6 (bilaga 1)

5.2.2 Modala satsadverbial på självständig språkanvändarnivå

Användningen av modala satsadverbial verkar öka betydligt mellan nivåerna A2 och B1 såväl i inlärtexterna som i läroböckerna. I inlärtexterna har andelen satsadverbial mångdubblats: den relativa frekvensen av satsadverbial på nivå B1 är 2 591,9 medan det på nivå A2 var bara 954,2. Att kunna säga hur signifikant den här stigningen i verkligheten är skulle dock förutsätta statistiska tester, vilka inom ramarna för den här avhandlingen inte är möjliga att utföra, men kunde vara värt att studera vidare i framtiden.

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	17	710,7
Ju	9	375,3
Naturligtvis	5	209,0
Tyvärr	4	167,2
Verkligen	3	125,4

Tabell 12. De fem mest frekventa modala satsadverbialen på nivå B1 i inlärtexterna.²⁵

De mest frekventa modala satsadverbialen i inlärtarmaterialet på nivå B1 presenteras i tabell 12. De utgörs av delvis samma ord som också var bland de vanligaste satsadverbialen i läroböckerna på lägre nivåer (*kanske*, *ju*, *verkligen*) men även ord som har varit mer sällsynta i läromaterialen förekommer (*naturligtvis*, *tyvärr*). Det allra mest frekventa modala satsadverbialet i inlärtarmaterialet på nivå B1 är *kanske* (rf=710,7), *ju* som följer på andra plats är mycket mer sällsynt (rf=375,3) och *naturligtvis*, *tyvärr* och *verkligen* är ännu mer marginella. De andra modala satsadverbialen som förekommer i materialet är *visst*, *gärna*, *förstås* och *enligt*, alla med ett enda belegg. Resten av orden som har taggats som satsadverbial i inlärtexterna på nivå B1 är inte modala satsadverbial. Att det ändå finns fem olika modala satsadverbial i materialet som används av minst två olika informanter visar att det inte är bara antalet satsadverbial som ökar när språkkunskaperna förbättras utan också variationen blir större – på nybörjarnivå var det ju bara *kanske* som systematiskt förekom i inlärtexterna.

²⁵ Sökning 4 (bilaga 1)

Satsadverbial	Absolut frekvens	Relativ frekvens
Ju	134	909,3
Kanske	117	794,0
Väl	48	325,7
Gärna	41	278,2
Verkligen	36	244,3

Tabell 13. De fem mest frekventa modala satsadverbialen på nivå B1 i COCTAILL.²⁶

En motsvarande sökning i COCTAILL-korpusen visar att användningen av modala satsadverbial ökar också i lärobokstexter på nivå B1, såsom framgår av tabell 13. Det är intressant hur *ju* igen är det mest frekventa modala satsadverbialet i läromaterialen, trots att det modala satsadverbialet som mest förekommer i inlärtexterna på samma nivå är *kanske*. Detta kan dock bero på att inlärnarna förväntas ha redan tillägnat sig användningen av ordet *kanske* och nu fokuserar läroboksförfattarna på att få dem att bli bekanta också med *ju*. Eftersom *ju* är ett mycket vanligt förekommande ord i L1-svenskan (se också avsnitt 5.4.3), är det nödvändigt att andraspråksinläraren kan åtminstone förstå ordet redan i ett tidigt skede av språkinläringen för att kunna kommunicera med modersmålstalare.

Det är överraskande hur orden *naturligtvis* och *tyvärr* är bland de mest frekventa satsadverbialen i inlärmaterialet på nivå B1 trots deras låga frekvenser i läromaterialen på samma nivå: *naturligtvis* med 12 träffar har en relativ frekvens på 81,4 och *tyvärr* med 11 träffar har en relativ frekvens på 74,7 i COCTAILL. Ordet *gärna* som däremot var det tredje mest frekventa modala satsadverbialet i läromaterialen på nivåerna A1 och A2 och den fjärde mest frekventa på nivå B1, förekommer bara en gång i en inlärttext på nivå B1.

Orsaker till att användningen av modala satsadverbial inte verkar gå hand i hand i inlärtexterna och läroböcker är närmast möjligt att spekulera om – en möjlig förklaring kan t.ex. ligga i uppgiftstypen där vissa typer av satsadverbial passar in bättre än andra. Att kunna dra sådana slutsatser skulle dock förutsätta en noggrann kvalitativ analys av

²⁶ Sökning 6 (bilaga 1)

såväl inlärtexterna som läromaterialen, vilket jag på grund av den här avhandlingens begränsningar tyvärr är tvungen att lämna till framtida forskning.

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	49	1084,5
Verkligen	17	376,3
Ju	11	243,5
Naturligtvis	8	177,1
Förstås/egentligen	6	132,8

Tabell 14. De fem mest frekventa modala satsadverbialen på nivå B2 i inlärtexterna.²⁷

De mest frekventa modala satsadverbialen i inlärtarmaterialet på nivå B2 presenteras i tabell 14. *Kanske* är igen det mest frekventa modala satsadverbialet i texterna, men ordet *verkligen* har trängt undan ordet *ju* som det andra mest frekventa modala satsadverbialet. *Ju* följer dock på tredje plats, men det är intressant hur låg dess relativa frekvens är på nivå B2 jämfört med nivå B1. En möjlig förklaring till detta är att uppsatserna på nivå B2 börjar vara mera formella och skriftspråkliga till stil, och därför undviker man sådana mer talspråkliga ord och uttryck som *ju*, som t.ex. enligt Josephson (2005: 5) är upp till tio gånger vanligare i tal än i skrift.

Att satsadverbialet *verkligen* är det andra mest frekventa modala satsadverbialet i inlärtexterna på nivå B2 är inte heller överraskande i sig eftersom ordet har varit ett av de mest frekventa modala satsadverbialen också i läromaterialen på lägre nivåer. Det är dock också helt möjligt att användningen av förstärkande modala satsadverbial reflekterar ett tilltagande självförtroende hos informanterna: känslan av att kunna uttrycka sig själv tillräckligt bra på främmande språk kan möjligtvis hjälpa språkinläraren att ha mod att uttrycka sig starkare.

Satsadverbial	Absolut frekvens	Relativ frekvens
Ju	123	760,5
Kanske	81	500,8
Väl	53	327,7
Gärna	48	296,8
Egentligen	29	179,3

Tabell 15. De fem mest frekventa modala satsadverbialen på nivå B2 i COCTAILL.²⁸

²⁷ Sökning 4 (bilaga 1)

²⁸ Sökning 6 (bilaga 1)

De fem mest frekventa modala satsadverbialen i läroböcker på nivå B2 är nästan samma som på nivå B1, såsom framgår av tabell 15. *Ju* och *kanske* är igen de klart mest frekventa modala satsadverbialen i materialet, och de används betydligt mycket mer än andra modala satsadverbial, men deras relativa frekvenser har gått ner en aning från de motsvarande talen på nivå B1 (*ju*: rf=909,3 → rf=760,5; *kanske*: rf=794,0 → rf=500,8). *Väl* och *gärna* följer på tredje och fjärde plats, såsom de gjorde också på nivå B1, och deras frekvenser är ungefär desamma som på nivå B1. På femte plats står däremot *egentligen* som inte tidigare har varit bland de fem mest frekventa modala satsadverbialen i läromaterialen, och som också knappt har förekommit i inläärartexterna, trots att ordet förekommer bland de vanligaste när man tittar på hela inläärarmaterialet. Det är dock värt att iakttä att såsom visades i tabell 8, har en stor del av träffarna av *egentligen* annoterats som någonting annat än satsadverbial i inläärarmaterialet och detsamma kan naturligtvis ha skett också med läroboksmaterialet.

5.2.3 Modala satsadverbial på avancerad nivå

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	53	1028,1
Ju	27	523,8
Egentligen	22	426,8
Enligt	20	388,0
Verkligen	18	349,2

Tabell 16. De fem mest frekventa modala satsadverbialen på nivå C1 i inläärartexter.²⁹

De fem mest vanligt förekommande modala satsadverbialen i inläärartexterna på nivå C1 presenteras i tabell 16. Det framgår av tabellen att de mest frekventa modala satsadverbialen på nivå C1 är ungefär desamma som på lägre nivåer, men bland dem finns ett nytt satsadverbial som inte förekommer bland de vanligaste på lägre nivåer, utan är faktiskt mycket sällsynt, nämligen ordet *enligt*. Trots att *enligt* är först det fjärde mest frekventa modala satsadverbialet i inläärartexterna på nivå C1 är det i ljuset av de relativa frekvenserna faktiskt ungefär lika frekvent som t.ex. ordet *verkligen* på nivå B2. I princip kunde det alltså vara motiverat att fråga om behärsknigen av ordet *enligt* kunde till och

²⁹ Sökning 4 (bilaga 1)

med ses som en indikator för att inläraren har nått nivån C1. En närmare granskning av träffarna visar dock en logisk förklaring till ordets höga frekvens. 90 % av träffarna förekommer i Tisus-korpusen, och eftersom en stor del av träffarna förekommer i en liknande textuell kontext (se ex. 35–37) blir det tydligt att studenterna i Tisus-testet verkar ha haft någon slags källtext som de har kunnat, kanske till och med förväntats att utnyttja och hänvisa till i uppsatsen, vilket naturligtvis syns i användningen av hänvisande satsadverbial. Det är även möjligt att konstruktionen *enligt X* har varit med i källtexten, vilket skulle naturligtvis ha gjort det lättare för informanterna att använda konstruktionen också själv.

- (35) **Enligt** Theorell kan stress ”skada vissa av hjärncellerna som reglerar minne och kortisonhalt i blodet”. (Student T57)
- (36) Människan är skapad för ett annat liv, **enligt** professor Theorell. (Student T62)
- (37) ”Stress är att ha för lite meningsfullt arbete”, **enligt** Georg Klein, professor vid Karolinska Institutet. (Student T66)

Det finns dock även fall där skribenten hänvisar till sin egen uppfattning om ämnet i stället för en utomstående källa. De här fallen utgör 25 % av träffarna:

- (38) Ett liv utan stress är **enligt** mig inte möjligt och skulle vara rätt tråkigt. (Student T15)
- (39) Ansvar är **enligt** mig en annan viktigt förklaring till att stressrelaterade besvär ökar i vårt samhälle. (Student T58)
- (40) Men ändå finns det **enligt** min uppfattning, åtminstone vissa allmänna regler som vanligtvis förenklar stresshanteringen. (Student T93)

Den höga frekvensen av ordet *enligt* kan alltså delvis förklaras med skrivuppgiftens typ, och därför är det på basis av det här materialet inte möjligt att dra slutsatser om den möjliga korrelationen mellan kunskapsnivån och användningen av ordet. Det kunde ändå vara befogat att i framtiden forska om kunskapen att överhuvudtaget hänvisa till olika källor är någonting som på riktigt etableras först när studenten har nått nivån C1.

Ordet *enligt* visar igen ett problem som uppstår i analysen om man bara tittar på de resultat som automatiskt genereras av korpusverktyget, eftersom *enligt* i sig inte är ett satsadverbial, utan den bestämmer alltid ett annat ord eller fras (se också avsnitt 2.1.6). Således borde t.ex. *enligt mig* och *enligt min uppfattning* räknas som två skilda modala satsadverbial, trots att de rent semantiskt uttrycker ungefär samma innehåll. I praktiken är det dock meningsfullt med tanke på den här undersökningens syfte att de olika

förekomsterna av konstruktionen *enligt X* räknas ihop för att se helhetsbilden av användningen av konstruktionen i fråga.

Ett annat märkvärdigt resultat är hur användningen av satsadverbial *ju* ökar igen i inlärttexter på nivå C1 efter att ha gått ner på nivå B2. Den nedgående användningen av *ju* kunde åtminstone delvis antas bero på att ju högre språkliga nivå som studenten når, desto mer formella och mindre vardagsspråkliga texter hen kan producera, vilket eventuellt kunde synas i lägre frekvenser av satsadverbial *ju*. Den ökande frekvensen på nivå C1 kan möjligtvis förklaras med ordets finare stilistiska nyanser som behärskas först på högre nivåer. Den här möjligheten diskuteras vidare i avsnitt 5.4.3.

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	92	541,0
Ju	65	382,3
Gärna	46	270,5
Verkligen	38	223,5
Egentligen	37	217,6

Tabell 17. De fem mest frekventa modala satsadverbialen på nivå C1 i COCTAILL.³⁰

I läroboksmaterialet visar användningen av modala satsadverbial en nedgående trend: på nivå B1 var den relativa frekvensen för satsadverbial 4363,4, medan det på nivå B2 var 3889,0. På nivå C1 är den relativa frekvensen för satsadverbial ännu lägre: 3787,3.³¹

På nivå C1 verkar frekvenserna av speciellt de sannolikhetsgraderande modala satsadverbialen ha blivit mindre: de två vanligaste satsadverbialen är *kanske* (rf=541) och *ju* (rf=382,3) som båda användes mer på nivå B2 (*kanske*: rf=760,5; *ju*: rf=500,8). *Gärna* som på nivå C1 är det tredje vanligaste modala satsadverbial i läroboksmaterialet användes ungefär lika mycket också på nivå B2 (rf=270,5 på C1, rf=296,8 på B2). Orden *verkligen* och *egentligen* används däremot mer på nivå C1 än på nivå B2 i läroboksmaterialet, vilket visar att distributionen mellan olika typer av modala satsadverbial verkar börja jämnas ut, och det är inte längre *kanske* och *ju* som klart dominerar användningen av modala satsadverbial i läroböcker.

³⁰ Sökning 6 (bilaga 1)

³¹ Sökning 8 (bilaga 1)

Satsadverbial	Absolut frekvens	Relativ frekvens
Kanske	10	3069,4
Tyvärr	2	613,9

Tabell 18. De mest frekventa modala satsadverbialen på nivå C2 i inlärtexter.³²

I mitt material finns det bara få inlärtexter som representerar den högsta kunskapsnivån, C2. Antagligen på grund av det låga antalet inlärtexter på nivå C2 finns det bara två modala satsadverbial som används av fler än en informant på den nivån. Ordet *kanske* förekommer totalt 10 gånger och ordet *tyvärr* bara två gånger i inlärtexterna på nivå C2, men på grund av det överhuvudtaget låga ordantalet i inlärtexter på den nivån är deras relativa frekvenser rätt höga: 3069,4 och 613,9 respektive. Förutom de här två orden förekommer även ordet *verkligen* två gånger som satsadverbial i materialet, men båda beläggen förekommer i en och samma text av en och samma informant. *Säkert, nog, naturligtvis, ju* och *förstås* förekommer också i inlärtexterna på nivå C2, alla med ett enda belägg. COCTAILL-korpusen täcker inte alls nivå C2.

På grund av det mycket lägre antalet ord på nivå C2 är den nivån inte direkt jämförbar med andra nivåer. Det är troligt att i ett större material skulle det återfinnas ett större antal modala satsadverbial på nivå C2, och att den lexikala variationen dessutom skulle vara större än på lägre nivåer, eftersom det på nivå C2 är fråga om väldigt avancerade språkkunskaper, och inlärares förväntas t.ex. att kunna ”skilja mellan finare nyanser även i mer komplexa situationer” (Skolverket, 2009: 24).

I det här kapitlet har jag granskat användningen av modala satsadverbial i inlärtexter och lärobokstexter. I följande avsnitt jämför jag resultaten vidare med ett oavhängigt förstaspråksmaterial, bloggtexter, för att se huruvida användningen av modala satsadverbial i inlärtexter reflekterar den faktiska språkanvändningen hos modersmålstalare, eller om det verkar förekomma över- eller underanvändning av något/några modala satsadverbial i inlärarespråket.

³² Sökning 4 (bilaga 1)

5.3 Modala satsadverbial i Bloggmix-materialet

För att jämföra hur användningen av modala satsadverbial skiljer sig mellan andraspråksinlärare och förstaspråkstalare, söker jag efter satsadverbial i Språkbankens 21 Bloggmix-korpusar. Sökningen ger totalt 4 370 644 träffar, och de fem mest frekventa modala satsadverbialen i bloggmaterialen presenteras i tabell 19.

Satsadverbial	Absolut frekvens	Relativ frekvens
Ju	1 361 060	2 210,7
Kanske	509 382	827,4
Verkligen	361 679	587,5
Väl	265 383	431,1
Gärna	188 507	306,2

Tabell 19. De fem mest frekventa modala satsadverbialen i Bloggmix-materialet.³³

Det framgår av tabellen att de fem mest frekventa modala satsadverbialen i bloggmaterialen är ungefär desamma som också förekommer bland de vanligaste i inlärmaterialet. Precis som i inlärmaterialet, är *ju* och *kanske* de två klart mest frekventa modala satsadverbialen också i Bloggmix-materialet, men andelen är rätt olika. Det finns ett påfallande avstånd mellan ordet *ju* och de följande satsadverbialen – antalet träffar på *ju* är nästan trefaldigt i förhållande till *kanske* som står på andra plats och som för sin del är relativt mycket mer frekvent i materialet än ordet *verkligen* som följer på tredje plats. Det är alltså tydligt att det allra mest frekventa modala satsadverbialen i den här typen av texter är *ju*.

I tabell 20 jämförs frekvenserna av de sju modala satsadverbialen, som förekom bland de vanligaste i inlärmaterialet och i bloggmaterialen, i studiens tre delmaterial.

³³ Sökning 9 (bilaga 1)

	L2-MATERIAL	L1-MATERIAL	
<i>Satsadverbial</i>	<i>SpIn, Tisus, Sw1203</i>	<i>COCTAILL</i>	<i>Bloggmix</i>
Kanske	898,2	508,3	827,4
Ju	305,8	561,8	2 210,7
Verkligen	254,8	163,3	587,5
Egentligen	178,4	139,4	235,0
Naturligtvis	165,6	87,3	47,3
Väl	51,0	197,1	431,1
Gärna	76,4	277,4	306,2

Tabell 20. Relativa frekvenserna av de mest frekventa modala satsadverbialen i inlärmaterialet och bloggmaterialet i studiens tre delmaterial.³⁴

Siffrorna visar att av de modala satsadverbialen som är mest frekventa i inlärmaterialet är det bara ordet *ju* som används betydligt mycket mera i bloggmaterialet, vilket förmodligen är ett tecken på en mer informell och talspråklig texttyp. I alla tre delmaterialen är *ju* dock ett mycket vanligt förekommande modalt satsadverbial. Frekvenserna av *kanske* är däremot överraskande lika i inlärtexter och i Bloggmix-materialet, vilket stödjer teorin om att den modersmålsliknande användningen av ordet *kanske* behärskas redan tidigt.

Ordet *naturligtvis* används väldigt lite i alla tre delmaterialen, men ändå är det dubbelt så frekvent i inlärmaterialet som det är i läromaterialen, och över tre gånger så vanligt i inlärmaterialet som i bloggmaterialet. Den här iakttagelsen ger anledning till närmare diskussion om en möjlig tendens hos andraspråksinlärare att överanvända ordet. Orden *väl* och *gärna* som däremot förekom bland de mest frekventa satsadverbialen i bloggmaterialet är mycket mindre representerade i inlärtexter. Den ringa användningen av ordet *väl* i inlärtexter kan inte förklaras med en låg frekvens av ordet i läromaterialen, eftersom ordet även i COCTAILL förekommer mer än t.ex. orden *verkligen*, *egentligen* och *naturligtvis*, som ändå alla är bland de mest frekventa satsadverbialen i inlärmaterialet. Ordet *gärna* är också överraskande sällsynt i inlärtexter trots att det har en relativt hög frekvens såväl i COCTAILL-korpusen som i Bloggmix-materialet, men det kan naturligtvis bero på att de evaluerande satsadverbialen inte överhuvudtaget

³⁴ Sökning 10 (bilaga 1)

är särskilt vanliga i inlärtexterna, till skillnad från satsadverbial för sannolikhet och försanthållande som även *väl* hör till.

Mer frekventa i inlärmaterialet	Mer frekventa i Bloggmix-materialet
<i>kanske; naturligtvis</i>	<i>ju; verkligen; egentligen; väl; gärna</i>

Tabell 21. Sammanfattande jämförelse av användningen av modala satsadverbial i L1- och L2-svenska.

Sammanfattningsvis kan det konstateras att när man tittar på hela inlärmaterialet används orden *kanske* och *naturligtvis* mer i inlärtexterna än i bloggtexter. Skillnaden är dock inte stor när det gäller ordet *kanske* vars användning faktiskt är överraskande lika i inlärmaterialet och Bloggmix-materialet. *Ju*, *verkligen*, *egentligen*, *väl* och *gärna* används däremot mer i Bloggmix-materialet. Den mest påfallande skillnaden är i användningen av orden *naturligtvis*, *väl* och *gärna*, vilket ger anledning till närmare granskning av resultat i avsnitt 5.4.1 för att se om det kan talas om en tendens hos andraspråksinlärare att över- eller underanvända de här tre orden.

		INLÄRARTEXTER				
Satsadvl.	Bloggmix	A2	B1	B2	C1	C2
Kanske	827,4	419,8 (-407,6)	710,7 (-116,7)	1084,5 (+257,1)	1028,1 (+200,7)	3069,4 (+2242)
Ju	2210,7	-	375,3 (-1835,4)	243,5 (-1967,2)	523,8 (-1686,9)	306,9* (-1903,8)
Verkligen	587,5	-	125,4 (-462,1)	376,3 (-211,2)	349,2 (-283,3)	613,9* (+26,4)
Egentligen	235,0	-	-	132,8 (-102,2)	426,8 (+191,8)	-
Naturligtvis	47,3	-	209,0 (+161,7)	177,1 (+129,8)	232,9 (+185,6)	-
Väl	431,1	-	-	66,4 (-364,7)	97,0 (-334,1)	-
Gärna	306,2	-	41,8* (-264,4)	66,4 (-239,8)	38,8* (-267,4)	-

Tabell 22. De relativa frekvenserna av modala satsadverbial i Bloggmix-materialet och i inlärtexterna på olika kunskapsnivåer (nivån där talet är närmast det motsvarande talet i Bloggmix i fetstil).³⁵

*Används av bara en informant.

³⁵ Sökning 11 (bilaga 1)

När det gäller olika kunskapsnivåer är det nivå B2 som verkar komma närmast den L1-enliga användningen av modala satsadverbial (tabell 22). En jämförelse mellan de sju modala satsadverbialen som är mest frekventa i hela inlärmaterialet och i Bloggmix-materialet visar att i tre av de sju fallen är frekvensen på nivå B2 närmast frekvensen i bloggmaterial. Det betyder dock inte att frekvenserna skulle i verkligheten komma särskilt nära varandra i bloggtexter och i inlärtexter på nivå B2 – t.ex. ordet *gärna* har en frekvens på 66,4 i inlärmaterialet på nivå B2, vilket är en mycket mindre frekvens än den motsvarande siffran i bloggmaterial (306,2). På nivåerna B1 och C1 är frekvensen ändå ännu lägre, och på nivåerna A2 och C2 förekommer ordet inte alls, vilket betyder att frekvensen på nivå B2 faktiskt kommer närmast frekvensen i bloggmaterial.

Att användningen av modala satsadverbial i L2-inlärtexter på nivå B2 verkar på basis av den här jämförelsen bäst motsvara den modersmålsliknande användningen tolkar jag att möjligtvis berättar om att språket och innehållet i inlärtexter på den nivån bäst motsvarar språket och innehållet i bloggtexter: ordförrådet och grammatisk kompetens på nivå B2 är tillräckligt utvecklade för att inläraren kan utan några större ansträngningar uttala sig om vardagliga ämnen, som ofta behandlas också i bloggar. På högre kunskapsnivåer kan såväl språket som innehållen vara mer formella och akademiskt inriktade, vilket för sin del inte motsvarar språket och innehållen i bloggtexter som ofta representerar ett mer informellt perspektiv.

5.4 Kvalitativ analys av resultaten

Trots att den här undersökningen först och främst är kvantitativ, finns det några aspekter i användningen av modala satsadverbial som jag vill diskutera närmare ur ett kvalitativt perspektiv.

5.4.1 Över- och underanvändningen av satsadverbial i inlärmaterialet

I avsnitt 5.3 kom det fram att det finns ett modalt satsadverbial som verkar användas mera i inlärtexter än i L1-materialen, och som alltså möjligen överanvänds av andraspråksinlärare. *Naturligtvis* har en relativ frekvens på 165,6 i hela inlärmaterialet,

medan de motsvarande talen är 87,3 i läroboksmaterialet och 47,3 i bloggmaterialiet. Frekvenserna är totalt sett inte höga, men det är ändå tydligt att ordet är mycket mer frekvent i inlärmaterialet jämfört med L1-materialen. Även när man tittar på enskilda kunskapsnivåer är resultaten desamma: ordet förekommer i inlärtexterna på nivåerna B1, B2 och C1 och på var och en av de här nivåerna verkar ordet vara överrepresenterat jämfört med bloggtexter. För att kunna fastslå hypotesen om överanvändningen av ordet *naturligtvis* hos andraspråksinlärare skulle inlärmaterialet dock behöva jämföras med motsvarande data av modersmålstalare för att kunna se om den flitigare användningen av ordet verkligen är ett för inlärarspråket utmärkande drag, eller om det helt enkelt bara är mer typiskt för inlärtexterna som texttyp än det är för bloggtexter.

Två ord som däremot dyker upp mycket i bloggmaterialiet men som är sällsynta såväl i hela inlärmaterialet som på enskilda kunskapsnivåer är *gärna* och *väl*. Båda orden används systematiskt även i lärobokstexter från och med A2, men de verkar ändå inte ha blivit etablerade i andraspråksinlärnarnas ordförråd, åtminstone inte på samma sätt som t.ex. *kanske*, *ju* och *verkligen*, som också är frekventa i läroböcker. Det är värt att fråga om de här två orden underanvänds av andraspråksinlärare. Det är inte en helt otvetydig slutsats. Ordet *gärna* anger evaluering – det betyder enligt Svensk Ordbok (2021) att någonting händer ”med nöje och med underförstått löfte om allmän positiv attityd”. Det är kanske just den evaluerande betydelsen som gör ordets användningsområde väldigt snävt – evaluerande modala satsadverbial verkar också annars vara väldigt sällsynta i inlärmaterialet. T.ex. ordet *tyvärr* förekommer i inlärmaterialet som satsadverbial totalt 15 gånger och dess relativa frekvens är 95,5.³⁶ Jämfört med t.ex. det negerande satsadverbial *inte* som har en absolut frekvens på 1 718 och en relativ frekvens på 10 943,5 i inlärmaterialet³⁷ är det tydligt att såväl *gärna* som *tyvärr* är mycket marginella i inlärtexterna. *Tyvärr* har dock relativt låga frekvenser också i L1-materialen (rf=42,2 i COCTAILL³⁸, rf=82,8 i Bloggmix³⁹), men det kan bero på att innehållen och den allmänna tonen i de här två texttyperna är snarare positiva. Att ordet *gärna* däremot används relativt mycket i lärobokstexter och bloggtexter, men inte i inlärtexterna, tolkar

³⁶ Sökning 12 (bilaga 1)

³⁷ Sökning 13 (bilaga 1)

³⁸ Sökning 14 (bilaga 1)

³⁹ Sökning 15 (bilaga 1)

jag kanske snarast som ett tecken på att det i läroböcker och bloggtexter helt enkelt finns mer utrymme för subjektiva evalueringar (se ex. 41 och ex. 42).

(41) Jag skulle **gärna** inreda en mysig gammal stuga vid havet med den här charmiga robusta stilen. (Bloggmix 2003)

(42) Jag vill att människor ska trivas i mitt sällskap och jag tar **gärna** ett glas i glada vänners lag. (COCTAILL)

Att ordet *väl* inte förekommer särskilt mycket i inlärtexter trots att det är mycket vanligt i L1-materialen, speciellt i bloggtexter, kan bero på att andraspråksinlärare underanvänder ordet, men det kan också bero på ordets funktioner som är delvis annorlunda i inlärtexter och i bloggtexter. Den här hypotesen kommer att diskuteras vidare i avsnitt 5.4.4.

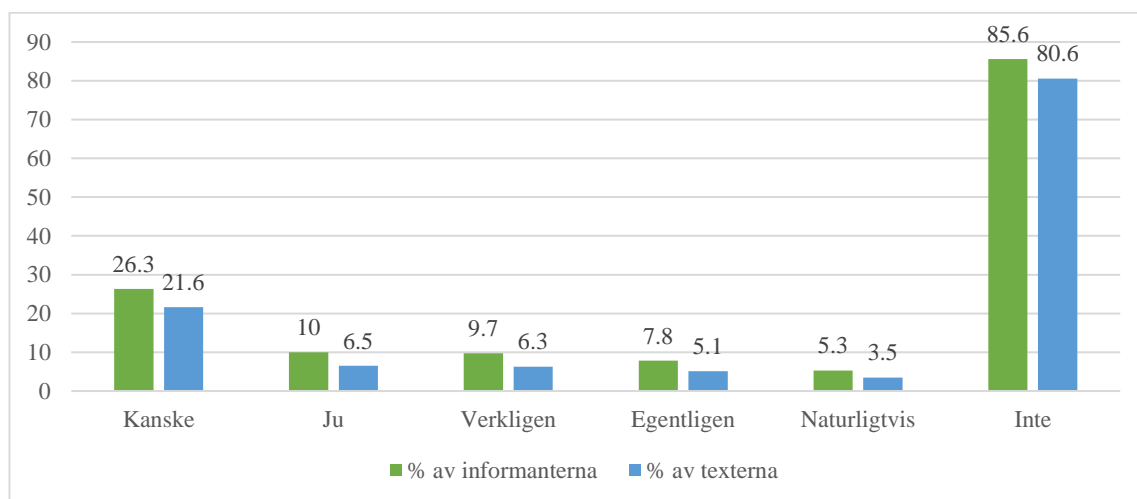
Något absolut svar på frågan om möjliga över- eller underanvändningstendenser av satsadverbial hos andraspråksinlärare är alltså inte möjligt att ge på basis av det här materialet, men användningen av orden *naturligtvis*, *gärna* och *väl* kunde vara värd att studera vidare för att se om deras användning i inlärarspråket verkligen skiljer sig från den målspråksenliga användningen.

5.4.2 Användningen av *kanske*

Vad som gör *kanske* ett speciellt intressant satsadverbial i ljuset av den här undersökningens resultat är att det används betydligt mycket mera än något annat modalt satsadverbial i inlärtarmaterialet, och det är det mest frekventa modala satsadverbial på alla kunskapsnivåer där modala satsadverbial överhuvudtaget förekommer. Dessutom är *kanske* det enda modala satsadverbial som förekommer i texterna redan på nybörjarnivå (A2). Det tycks således vara det allra första modala satsadverbial som tillägnas av andraspråksinlärare. Ordet är tydligen rätt speciellt, och därför vill jag analysera dess användning närmare.

Som sagt, används ordet *kanske* mycket mer än något annat modalt satsadverbial i materialet. Ordet förekommer som satsadverbial i inlärtarmaterialet totalt 141 gånger och det används i totalt 106 olika texter. Eftersom några informanter har skrivit två eller tre skilda texter blir antalet enskilda informanter som använder ordet *kanske* sammanlagt 84. Det betyder att ca 26 % av informanterna i inlärtarmaterialet använder ordet minst en gång, vilket är mycket jämfört med de andra modala satsadverbialen i materialet, såsom framgår

av figur 2 som visar hur användningen av modala satsadverbial har distribuerats i materialet, dvs. hur stor andel av enskilda informanterna som använder ordet och hur stor andel av enskilda texterna som innehåller ordet.



Figur 2. Andelen av informanter som använder orden *kanske*, *ju*, *verkligen*, *egentligen*, *naturligtvis* och *inte* i sin text samt andelen av inlärtartexter som innehåller de här orden.⁴⁰

	A1	A2	B1	B2	C1	C2
Kanske	0 %	9 %	20 %	38 %	41 %	80 %
Ju	0 %	0 %	9 %	9 %	20 %	20 %
Verkligen	0 %	0 %	5 %	15 %	19 %	20 %
Egentligen	0 %	0 %	0 %	8 %	21 %	0 %
Naturligtvis	0 %	0 %	5 %	9 %	10 %	20 %
Inte	44 %	73 %	94 %	95 %	99 %	100 %

Tabell 23. Den procentuella andelen av informanter som använder orden *kanske*, *ju*, *verkligen*, *egentligen*, *naturligtvis* och *inte* minst en gång i sin text.⁴¹

Av enskilda kunskapsnivåer är användningen av *kanske* som högst på nivå C2 där 80 % av informanterna använder ordet. C2 är dock inte jämförbar med de andra nivåerna på grund av det låga antalet texter på den nivån. Även på nivåerna B1, B2 och C1 används ordet av en större andel av informanterna än något annat modalt satsadverbial – på nivå B1 används *kanske* av en femtedel av informanterna, på nivå B2 ca 38 % av informanterna

⁴⁰ Sökning 16; sökning 17 (bilaga 1)

⁴¹ Sökning 18 (bilaga 1)

och på nivå C1 av 41 % av informanterna. Jämfört med ordet *inte* är andelarna dock inte höga.

Att satsadverbialen *kanske* är så mycket mera frekvent i inlärttexter än något annat modalt satsadverbial kan bero på flera faktorer. En möjlig förklaring ligger i ordets semantik. Ordet *kanske* betyder enligt Svensk Ordbok (2021) att någonting händer ”med viss möjlighet”. Det är ett stilistiskt neutralt sätt att uttrycka att någonting inte är garanterat: genom att använda ordet *kanske* behöver man inte ta närmare ställning till huruvida sakförhållandet är mera sannolikt eller osannolikt. Att det redan i början av språkinlärningsprocessen finns ett behov för ett ord som uttrycker att någonting inte är garanterat beror för sin del på de underförstådda normer som styr vår kommunikation. En grundläggande förutsättning för fungerande kommunikation är att alla samtalsparter samarbetar. I praktiken kan alla samtalsdeltagare förutsättas följa de så kallade Grices samtalsmaximer (*conversational maxims*) för att skapa ett rationellt och kooperativt samtal. En av Grices samtalsmaximer hänför sig till samtalskvalitet: talaren ska inte uttrycka som sanning sådant som inte nödvändigtvis är sant. På lyssnarens sida gäller det att hen ska förmoda att det som hen hör är sant om det inte finns bevis på annat. (Saeed, 2016: 211.) Därför är det ytterst viktigt att språkinläraren redan på nybörjarnivå kan göra lyssnaren medveten om de möjliga restriktionerna i yttrandets sanningshalt. Det konkreta behovet för ett enkelt och entydigt ord som uttrycker att någonting inte är garanterat och som kan användas oberoende av hur sannolikt sakförhållandet kan estimeras vara kan förklara varför *kanske* av alla modala satsadverbial verkar tillägnas först av andraspråksinlärare, och varför det också senare i språkinlärningsprocessen används allra mest.

Kanske hör till de sannolikhetsgraderande modala satsadverbialen, men sannolikhetsgradering är inte dess enda funktion. Nilsson (2005: 141) märkte att ordet *kanske* kan även fungera bl.a. i förmildrande funktion och som markör för exemplifierande urval, samt i spekulerande kontexter.

I förmildrande funktion används *kanske* för att tona ner ett sakpåstående (Nilsson, 2005: 141). Den här funktionen av *kanske* förekommer också i mitt material, såsom framgår av exemplen 43–45.

(43) Det vanligaste levnadssätt idag är **kanske** att vara sambo. (Student W8)

(44) Stresset är **kanske** ett nytt fenomen i Sverige. (Student T58)

- (45) Det går **kanske** inte att göra så mycket åt den stress som är framkallad av arbetslösheten.
(Student T68)

I exemplen använder skribenterna ordet *kanske* för att förmildra sakpåståenden som annars borde tolkas som sanna. Valet av att använda ett förmildrande satsadverbial kan bero på tveksamheten över sitt eget kunnande, men kan också reflektera studenternas självförtroende över sina egna språkfärdigheter. Att uttrycka sig själv på ett främmande språk kan vara skrämmande och speciellt i en stressfull provsituation kan inläraren känna sig osäker. I fall studenten inte litar på sin egen kunskap om ämnet eller är i övrigt tveksam om att uttrycka sina tankar som sanningar, kan hen välja att mildra informationsinnehållet med ordet *kanske*. Den motsatta reaktionen, dvs. att satsadverbial används för att förstärka egna argument, uttrycks i mitt material bl.a. med ordet *verkligen*. Användningen av förstärkande satsadverbial kan vittna om en bättre självtroende när det gäller inlärarens egna språkfärdigheter och övriga kunskaper (se ex. 46–48). Den här hypotesen stöds även av det faktumet att användningen av ordet *verkligen* ökar i takt med att språkkunskaperna i övrigt förbättras: på nivå B1 är ordets relativa frekvens 125,4 medan på nivå B2 har frekvensen trefaldigats (rf=376,3). På nivå C1 kommer frekvensen ner en aning (rf= 349,2) men ordet är fortfarande nästan tre gånger lika frekvent som på nivå B1.

- (46) Man **verkligen** arbetar endast för att kunna ha råd att leva. (Student T29)
(47) Man skulle kunna jobba så många timmar som man gör nu utan man blir arbetskadad men då skall man **verkligen** tycka om det man arbetar med. (Student T55)
(48) Men skulle man **verkligen** gå tillbaka till samma arbete när man blev sjuk från jobbet?
(Student W48)

I Nilssons material används ordet *kanske* även för att uttrycka exemplifierande urval, vilket enligt Nilsson uttrycker att ”det bestämda ledet är ett exempel på en av flera möjliga referenter som propositionen gäller för” (2005: 71). Sådan användning av ordet *kanske* förekommer också i mitt material, såsom i exempel 49 där skribenten använder ordet för att uttrycka att bland de saker som gör hen lycklig är fåglar och blommor, men det är tydligt att de är exempel på flera olika möjliga referenter:

- (49) Jag kan se **kanske** en fågel eller en vacker blomma och jag bli fyllt av lycka och lugn.
(Student W50)

Ett tredje användningsområde som Nilsson fann för ordet *kanske* och som också förekommer i mitt material är spekulationer. *Kanske* används då för att markera att det som sägs är spekulation om hur ”saker och ting förhåller sig eller kommer att förhålla sig i framtiden” (Nilsson, 2005: 202). Den här användningen verkar också vara rätt vanligt i mitt material, vilket kan dock vara kännetecknande för texttypen.

(50) han ville lära sig dansa balett men sin papa vill inte, **kanske** han tanker balett är bara för tjejer och om du är pojke, du kan inte göra. (Student S70)

(51) Det var **kanske** så för sextio, sjuttio år sedan då skillsmässor fördömdes av omgivningen. (Student W8)

(52) Då **kanske** kommer också mandarinkinesiska att dominera inom olika områden. (Student W34)

Den stora frekvensen av ordet *kanske* kan alltså förklaras med ordets användbarhet och dess breda användningsområde. I nästa avsnitt diskuterar jag användningen av ordet *ju* som också är ett mycket vanligt förekommande satsadverbial i hela forskningsmaterialet och således värt att titta närmare på.

5.4.3 Användningen av *ju*

Ju är ett vanligt förekommande modalt satsadverbial såväl i inlärttexter som i L1-materialen. Om man bara tittar på de belägg där ordet har taggats som satsadverbial är *ju* det mest frekventa modala satsadverbial i COCTAILL-korpusen och i Bloggmix-materialet, och det andra mest frekventa i hela inlärtmaterialet. Användningen av ordet börjar först på nivå B1 i inlärttexterna, men i läroboksmaterialet är *ju* faktiskt det mest frekventa modala satsadverbial redan på nivå A1, och på nivå A2 har ordet *kanske* bara knappt högre frekvens än *ju*. På nivåerna B1 och B2 är *ju* klart det vanligaste modala satsadverbial i läroböcker.

Den stora frekvensen av ordet *ju* i lärobokstexter kan möjligtvis förklaras med hur frekvent ordet är i L1-svenskan. En sökning efter satsadverbial i Bloggmix-materialet visade att *ju* används många gånger mer än något annat modalt satsadverbial i bloggtexter, och en snabb jämförelse med andra förstaspråkskorpusar i Korp visar att *ju* också är det mest vanligt förekommande modala satsadverbial t.ex. i skönlitterära texter,⁴² Twitter-

⁴² Sökning 19 (bilaga 1)

texter⁴³ och tidningstexter,⁴⁴ och kanske överraskande nog, även i akademiska texter,⁴⁵ dock med en relativt låg frekvens (rf=346,9). Dessutom har ordet noterats vara mycket frekvent som diskurspartikel (se t.ex. Svensson, 2009). Därför kan läroboksförfattare ha velat introducera andraspråksinlärare till ordet så tidigt som möjligt – för att stödja dem i kommunikationen med modersmålstalare.

Att *ju* förekommer bland de vanligaste satsadverbialen också i inlärartexter kan således bero på den stora exponeringen för ordet i lärobokstexter. Det kan ändå också kopplas till begreppen *formell* och *informell inläring* som diskuterades i kapitel 2.2.3 – att det inte bara är den explicita, undervisningsbaserade inläringen i klassrumsmiljö som påverkar det faktiska språkanvändningen hos andraspråksinlärare, utan även den språkliga exponeringen utanför klassrummet har betydelse till det. Med detta menar jag inte bara den direkta kommunikationen med modersmålstalare (t.ex. samtal) utan även det språkliga inflytandet via olika medier, såsom bl.a. tv, radio och internet. Också Viberg (2004: 197) påpekar att en stor del av ordinläringen faktiskt sker *incidentiellt* – dvs. som en biprodukt när man t.ex. deltar i samtal eller tittar på tv och försöker förstå vad som sägs. Att ordet *ju* används i stor utsträckning i inlärartexter beror förmodligen på den stora exponeringen för ordet såväl inne i klassrummet som utanför det, och inläringen av ordet kan således ha skett både explicit och implicit.

En intressant iakttagelse angående användningen av ordet *ju* i inlärmaterialet var hur användningen av ordet minskade mellan nivåerna B1 (rf= 375,3) och B2 (rf= 243,5) men ökade igen på nivå C1 (rf= 523,8). Sänkningen i användningen av *ju* kan förmodligen åtminstone delvis förklaras med hur talspråkigheter stryks när tonen i inlärartexter också i övrigt blir mer formell, men vad den branta stigningen i användningen av ordet mellan nivåerna B2 och C1 beror på är en mer komplex fråga.

Enligt Svensk ordbok (2021) står ordet *ju* för någonting som kan ”betraktas som bekant”. Den korta definitionen gör inte rättvisa för det stora antalet funktioner som ordet kan presentera i satsen. Josephson (2005: 9) nämner bl.a. användningen av ordet *ju* som samförståndsord för att visa att både sändaren och mottagaren är införstådda, användningen i nedtonande funktion och användningen i syfte att förklara för mottagaren

⁴³ Sökning 20 (bilaga 1)

⁴⁴ Sökning 21 (bilaga 1)

⁴⁵ Sökning 22 (bilaga 1)

vilken information i yttrandet som är eller borde vara känt och vilken är ny. Ordet kan dessutom användas som övertalningsord i meningsutbyten, vilket Josephson exemplifierar på följande sätt:

Ju låter förstå att talaren har goda skäl för det som sägs. ”Sverige har ju halkat efter i välfärdslihan” eller ”det blir ju ingen riktig jul utan Örebrokusinerna” kan vara mycket kontroversiella påståenden. Med ett *ju* låter de nästan självklara. I debattsituationer tycks *ju* vara särskilt vanligt i invändningar, ”men du har ju fel” och när talaren omöjligen kan ha fullständig kännedom om det som sägs, ”vi kommer ju alla bli fattigare om Sverige står utanför EMU.” (Josephson, 2005: 10.)

Ju kan alltså uttrycka ett stort antal olika funktioner, varav några tydligen finare än andra. Även Tandefelt (2005) beskriver ett välplacerat *ju* som ett ”skarpt retoriskt vapen”. Genom att använda *ju* på mångsidigt och kreativt sätt kan man alltså visa att man känner till de finaste nyanserna som *ju* kan bära med sig till satsen, vilket för sin del kan berätta om en avancerad språkförmåga och en högutvecklad språkkänsla.

I mitt material används ordet *ju* i flera olika funktioner, t.ex. förmildrande (ex. 53), i samband med kontroversiella påståenden (ex. 54), i invändningar (ex. 55) och även i yttranden där det faktiskt är tydligt att skribenten inte kan ha fullständig kännedom om det som hen påstår (ex. 56). Användningen av ordet helt enkelt i betydelsen ”kan betraktas som bekant” förekommer också, såsom framgår av exemplet 57 där skribenten beskriver handlingen i en bok.

- (53) Vi har också brister, vi är **ju** människor, inga robotar man kan bara uppdatera eller renovera från små bitar. (Student T77)
- (54) En liten arbetare kan klaga sina rättighets problem vidare men det hjälper **ju** inte, faktiskt. (Student T29)
- (55) Ännu längre arbetsdagar som pågår hemma har gjort våra liv arbetsorienterade och vi hinner inte njuta av fritiden – det finns **ju** ingen. (Student T67)
- (56) För länge sen kunde man **ju** bo i vilken grotta som helst, man kunde **ju** äta från vilken träd man ville och vilken djur man kunde fånga. (Student T103)
- (57) Arne visste **ju** att hon hade ett barn. (Student S55)

Sammanfattningsvis kan den höga frekvensen av ordet *ju* i inlärttexter alltså antas bero på dess höga frekvens i L1-svenskan. Att ordet är så vanligt förekommande i tal och texter av förstaspråkstalare kan å ena sidan ha bidragit till valet av att använda ordet mycket också i läroböcker, vilket för sin del har lett till att informanterna har fått mycket

exponering för ordet redan i början av språkinlärningsprocessen. Å andra sidan är det troligt att informanterna också utanför klassrummet har stött på ordet ofta när de t.ex. har tittat på tv, surfat på sociala medier eller umgåtts med kompisar, och så småningom har ordet blivit befäst i inlärares språksystem. Att *ju* används mycket också på högre kunskapsnivåer kan möjligen förklaras med de finaste nyanserna som *ju* kan uttrycka, trots att ordet är särskilt vanligt i talspråk och således skulle den intuitiva hypotesen ha varit att ju högre kunskapsnivå, desto mindre borde *ju* faktiskt synas i texter, eftersom den allmänna tonen och stilen i texterna kunde antas vara mer formella och skriftspråkliga. På basis av det här materialet är det här dock inte fallet, utan *ju* används flitigt också på högre kunskapsnivåer.

5.4.4 Användningen av *väl*

Ett sista modalt satsadverbial vars användning jag vill behandla närmare är *väl*. I avsnitt 5.3 kom det fram att ordet *väl* är mycket mer frekvent i lärobokstexter (rf=197,1) och i bloggtexter (rf=431,1) än i L2-inlärtexter (rf= 51,0 i hela inlärtarmaterialet). Den här iakttagelsen väckte frågan om en möjlig tendens hos andraspråksinlärare att underanvända ordet, men det är inte en entydig slutsats. För att kunna analysera användningen av ordet närmare, är det värt att utgå från dess definition.

Med ordet *väl* kan talaren ange att ”han försöksvis antar det som sägs i en deklarativ huvudsats med påståendefunktion, i en hävdad bisats eller i en bisats som utan *väl* skulle uppfattas som presupponerat sann” (Teleman m.fl. 1999b: 116). *Väl* kan även användas för att göra deklarativa huvudsatser till försiktiga frågor, och i sådana satser som uttrycker t.ex. värderingar, avsikt eller tvång, kan talaren med *väl* ange att hen ansluter sig till värderingen och väntar på mottagarens medhåll (ibid.). Svensk ordbok (2021) ger ordet tre definitioner: ”som man måste förmoda”, ”efter de kända omständigheterna att döma” samt ”i vädjande uttryck”. Ordet kan alltså beskrivas ha en sannolikhetsgraderande ton, men det kan även användas för att uttrycka att talaren eller skribenten väntar på någon slags reaktion hos lyssnaren. Det är just de här interaktionella funktionerna som kan å ena sidan förklara den höga frekvensen av *väl* i bloggtexter och å andra sidan den låga frekvensen i inlärtexter – medan inlärtexter kan beskrivas som monologisk texttyp, kan bloggtexter ses som snarare dialogiska (se också avsnitt 4.2.2).

(58) Nu börjar man faktiskt känna sig lite rund om ansiktet, det brukar **väl** vara så i slutet av graviditeten? (Bloggmix 2005)

(59) Ni minns **väl** topparna som jag beställde från Zara? (Bloggmix 2011)

I exemplen 58 och 59 vädjar bloggskribenten till läsaren att få någon slags reaktion. I exempel 58 gör skribenten en deklarativ huvudsats till en försiktig fråga med ordet *väl*, och på detta sätt vänder sig till läsaren för att få en bekräftelse för sin uppfattning. Exemplet 59 beskriver tydligt den dolda dialogen som finns mellan skribenten och läsaren, och som har funnits också tidigare. Trots att det ytligt sett är skribenten som är den aktiva deltagaren i samtalet mellan sändaren och mottagaren, är det faktiskt fråga om ett slags samtal mellan dessa två. *Väl* hjälper alltså skapa intersubjektivitet mellan skribenten och mottagaren, vilket enligt Lindström (2008: 96) är en viktig funktion hos modalpartiklar som även *väl* representerar.

I inlärmaterialet används ordet *väl* enbart för att uttrycka att det som skrivs är förmodan:

(60) Men viktigast är **väl** att de blir friska igen och inte får problem med alkohol eller andra droger. (Student W48)

(61) Det finns alltså mindre jobb och även om man har jobb är det **väl** möjligt att man snabbt kan bli arbetslös. (Student T56)

Den samtalsskapande funktionen som förekommer med ordet *väl* i bloggtexter saknas i inlärttexter, vilket som sagt beror på att inlärttexter som texttyp är snarare monologiska och väcker inte ömsesidig kommunikation på samma sätt som bloggtexter. En bloggtext kan kännas som ett samtal mellan två kompisar, skribenten och läsaren, trots att relationen mellan dessa två ofta är ensidig: läsaren vet vem skribenten är, men för skribenten är läsaren ofta bara en anonym medlem i läsarkretsen (se t.ex. Brake, 2009). När det gäller inlärttexter kan skribenten och textens mottagare, t.ex. läraren, känna varandra personligt och relationen mellan dessa två kan således vara mer förtrolig än relationen mellan bloggskribenten och läsaren.⁴⁶ Trots det kan en bloggtext vara ett mera förtroligt och informellt samtal mellan parterna, vilket för sin del kan realiseras bl.a. i användningen av modala satsadverbial.

⁴⁶ I mitt material är det dock inte lika entydigt att säga vem som den tilltänkta läsaren har varit när inlärttexterna har skrivits, eftersom texterna är insamlade i olika typer av sammanhang. T.ex. i Tisus kan den tilltänkta läsaren ha varit en anonym testbedömare, medan i Sw1203 kan informanterna ha tänkt på kursens lärare som mottagare till texterna.

På basis av den här undersökningen är det alltså inte möjligt att konstatera att ordet *väl* skulle underanvändas i inlärttexter, trots att det baserat på frekvenserna verkar vara så. Ordet används tydligen för att uttrycka fler funktioner i bloggtexter, och således kan den flitigare användningen av *väl* möjligtvis förklaras med ordets polysemiska karaktär som bättre behärskas och utnyttjas av förstaspråkstalare.

5.5 Annoteringens påverkan på resultaten

Ett av avhandlingens syften är att ta reda på hur mycket den automatiska annoteringen påverkar studiens resultat, och på basis av den här kunskapen evaluera hur automatiskt annoterade korpusar överhuvudtaget lämpar sig för den här typen av undersökning. För att kunna utvärdera materialvalets lämplighet är det nödvändigt att granska hur lyckad annoteringen i själva verket är.

En närmare titt på träffarna av ord och uttryck som har taggats som satsadverbial i inlärtarmaterialet visar att en stor del av dem faktiskt inte är modala satsadverbial, i några fall inte ens satsadverbial, eller i värsta fall, inte ens adverbial. I exemplen 62–64 visas ord som har taggats som satsadverbial trots att de inte är det:

(62) Då åt vi middag och **pratade** om reser. (Student S59)

→ Annoterats som satsadverbial, är predikat.

(63) Ju mer man vet, desto mer **kräver** man. (Student T103)

→ Annoterats som satsadverbial, är predikat.

(64) **Igår** vaknade jag ha stressad. (Student S42)

→ Annoterats som satsadverbial, är tidsadverbial.

Det är klart att de felaktigt taggade orden har förvridit resultaten. T.ex. att hävda att det totala antalet satsadverbial i inlärtarmaterialet skulle vara 555, såsom räknas av korpusverktyget, är en slutsats som man bör förhålla sig kritiskt till. Ett annat konkret exempel på hur de ord som taggats som satsadverbial, men som i själva verket inte är det, förvränger resultat är att det t.ex. i inlärttexter på kunskapsnivån A1 inte förekommer några modala satsadverbial, trots att korpusverktyget hittar totalt 3 träffar för satsadverbial i den delen av inlärtarmaterialet. Tre felaktiga belägg låter kanske inte mycket, men normaliserad till 1 000 000 ord, blir den relativa frekvensen 580, trots att

det faktiskt borde vara 0. Här syns dock tydligt också problematiken med att normalisera till 1 000 000 ord med en mindre korpus – även ett enda felaktigt belägg antar större proportioner än vad som kanske skulle vara nödvändigt.

Resultaten påverkas inte bara av ord och uttryck som felaktigt taggats som satsadverbial, utan även av ord som är satsadverbial men som har taggats som någonting annat. Det här problemet behandlades redan ytligt i avsnitt 5.1 där det noterades att de fem modala satsadverbialen som används mest i inlärmaterialet, förekommer faktiskt fler gånger i materialet än vad som har taggats som satsadverbial. Det konstaterades att orsaken till detta kan självklart vara att de helt enkelt förutom satsadverbial också förekommer som andra satsdelar, men likaväl är det möjligt att de belägg som inte har taggats som satsadverbial i själva verket är satsadverbial, och således borde de ingå i resultaten.

Inom ramarna för den här undersökningen är det inte möjligt att gå igenom varje möjligt satsadverbial i materialet och räkna hur många belägg som inte har taggats som satsadverbial för att evaluera hur resultaten i verkligheten borde se ut. Därför har jag valt att närmare granska de fem mest frekventa modala satsadverbialen i inlärmaterialet, vilka som redan noterades, förekommer fler gånger i materialet än vad som har annoterats som satsadverbial. Jag utreder vilka andra dependensrelationer som de här orden har fått i materialet och jämför resultatet med en manuellt annoterad kontrollkorpus, Talbanken, som också är tillgänglig i Korp. Resultaten av den här jämförelsen har samlats in i tabell 24.

<i>Satsadverbial</i>	<i>Dependensrelation</i>	<i>Talbanken</i>	<i>Inlärmaterialet</i>
<i>Kanske</i>	Satsadverbial	93 % (70/75)	79 % (141/178)
	Koordination på huvudsatsnivå	3 % (2/75)	0 % (0/178)
	Prepositions komplement	1 % (1/75)	1 % (2/178)
	Makrosyntagm	1 % (1/75)	0,5 % (1/178)
	Annat adverbial	1 % (1/75)	4 % (8/178)
	Kontrastivt adverbial	0 % (0/75)	7 % (13/178)
	Huvud	0 % (0/75)	1 % (2/178)
	Konjunktionellt adverb	0 % (0/75)	1 % (2/178)
	Framförställt attribut	0 % (0/75)	1 % (2/178)
	Rot	0 % (0/75)	1 % (2/178)
	Efterställd bestämning	0 % (0/75)	0,5 % (1/178)

	Determinerare, bestämningsord	0 % (0/75)	0,5 % (1/178)
	Samordnat led	0 % (0/75)	0,5 % (1/178)
	Tidsadverbial	0 % (0/75)	0,5 % (1/178)
	Subjektspredikativ (subjektiv predikatsfyllnad)	0 % (0/75)	0,5 % (1/178)
<i>Ju</i>	Satsadverbial	100 % (34/34)	96 % (48/50)
	Annat adverbial	0 % (0/34)	2 % (1/50)
	Kontrastivt adverbial	0 % (0/34)	2 % (1/50)
<i>Verkligen</i>	Satsadverbial	85 % (11/13)	67 % (40/60)
	Annat adverbial	15 % (2/13)	23 % (14/60)
	Determinerare, bestämningsord	0 % (0/13)	3 % (2/60)
	Rot	0 % (0/13)	2 % (1/60)
	Negerande adverbial	0 % (0/13)	2 % (1/60)
	Kontrastivt adverbial	0 % (0/13)	2 % (1/60)
	Framförställt attribut	0 % (0/13)	2 % (1/60)
<i>Egentligen</i>	Satsadverbial	100 % (12/12)	43 % (28/65)
	Annat adverbial	0 % (0/12)	40 % (26/65)
	Kontrastivt adverbial	0 % (0/12)	5 % (3/65)
	Efterställd bestämning	0 % (0/12)	3 % (2/65)
	Komparativt adverbial	0 % (0/12)	2 % (1/65)
	Huvud	0 % (0/12)	2 % (1/65)
	Oklassificerbar satsfunktion	0 % (0/65)	2 % (1/65)
	Underordnad sats (bisats), utom subjunktion	0 % (0/65)	2 % (1/65)
	Subjektspredikativ (subjektiv predikatsfyllnad)	0 % (0/65)	2 % (1/65)
	Samordnat led	0 % (0/65)	2 % (1/65)
<i>Naturligtvis</i>	Satsadverbial	91 % (21/23)	87 % (26/30)
	Koordination på huvudsatsnivå	4 % (1/23)	0 % (0/30)
	Annat adverbial	4 % (1/23)	7 % (2/30)
	Rot	0 % (0/23)	3 % (1/30)
	Efterställd bestämning	0 % (0/23)	3 % (1/30)

Tabell 24. Modala satsadverbials dependensrelationer i inläromaterialet och i Talbanken.⁴⁷ Dependensrelationerna har skrivits i den form som de står i Korp.

⁴⁷ Sökning 23 (bilaga 1)

Tabellen visar att alla de fem modala satsadverbialen har fått minst tre olika kategorier i inlärmaterialet, medan i Talbanken är det bara ordet *kanske* som överhuvudtaget har fått över tre kategorier. Procentuellt har de här fem orden klassificerats som satsadverbial i 85–100 % av fallen i Talbanken, medan i inlärmaterialet varierar den procentuella andelen av satsadverbial på en större skala, från 43 % till 96 %. Skillnaden är överraskande stor och ger anledning till en närmare kvalitativ analys.

Ordet *kanske* har det största antalet träffar såväl i inlärmaterialet som i Talbanken, och det har också fått det största antalet olika kategorier i båda materialen. I inlärmaterialet finns det 14 olika kategorier för *kanske*, varav satsadverbial utgör en tydlig majoritet av beläggen (79 %). En titt på de belägg som inte har taggats som satsadverbial visar att alla de 37 träffarna kunde även ha taggats som satsadverbial. Det betyder att det totala antalet träffarna av *kanske* som satsadverbial skulle i stället för 141 vara 178. Exempelen 65–67 presenterar några felaktigt taggade belägg av ordet *kanske*:

(65) Det var **kanske** så för sextio, sjuttio år sedan då skillsmässor fördömdes av omgivningen.

(Student W8)

→ Taggats som annat adverbial (tagget AA), är satsadverbial (tagget MA)

(66) Så varför vill de inte gifta sig istället: **kanske** för att de är rädd att de inte kan hålla ihop för en lång tid. (Student W18)

→ Taggats som kontrastivt adverbial (CA), är satsadverbial (MA)

(67) Ibland **kanske** lösningen är att få vara hemma och ladda batterierna ett tag. (Student T87)

→ Taggats som subjektspredikativ (SP), är satsadverbial (MA)

Ju har kategoriserats som satsadverbial i 48 av de totala 50 beläggen i inlärmaterialet.⁴⁸ I Talbanken har alla förekomster av *ju* taggats som satsadverbial. Av de två fallen där ordet har taggats som någonting annat än satsadverbial i inlärmaterialet borde ett tydligen ha taggats som satsadverbial (ex. 68). I det andra belägget är satsinnehållet oklart, och det är svårt att säga vad skribenten har åsyftat att skriva. Det totala antalet träffar av *ju* som satsadverbial i inlärmaterialet borde således vara 49.

⁴⁸ En sökning efter *ju* i inlärmaterialet ger 54 träffar, men fyra av träffarna är förekomster av konjunktionen *ju-desto* som jag har räknat bort från resultaten.

(68) För länge sen kunde man **ju** bo i vilket grotta som helst, man kunde ju äta från vilken träd man ville och vilken djur man kunde fånga. (Student T103)

→ Kategoriserats som annat adverbial (AA), är satsadverbial (MA)

Ordet *verkligen* förekommer totalt 60 gånger i inlärmaterialet, men en tredjedel av beläggen är inte annoterade som satsadverbial. I Talbanken har 85 % av förekomsterna av *verkligen* annoterats som satsadverbial. En närmare granskning av träffarna visar att av de 20 fallen där *verkligen* har taggats som någonting annat än satsadverbial i inlärmaterialet kunde 19 faktiskt ha kategoriserats som satsadverbial. Det totala antalet träffar på ordet *verkligen* som satsadverbial borde alltså vara 59. Det enda fallet där ordet inte kan räknas som modalt satsadverbial är en mening där informanten förmodligen av misstag blandar ihop adverbet *verkligen* och adjektivet *verklig* (ex. 69). *Verkligen* i det här fallet har annoterats som *framförställt attribut*.

(69) Det är svår att ha på egen hand en **verkligen** idé från världen. (Student W3)

Den mest påfallande skillnaden mellan annoteringarna i Talbanken och inlärmaterialet är med ordet *egentligen*. I Talbanken har varje träff av ordet taggats som satsadverbial, medan i inlärmaterialet har bara 28 av de totala 65 träffarna hamnat under kategorin *satsadverbial*. Det betyder att över hälften av träffarna har taggats med andra kategorier. En titt på träffarna visar ändå att vartenda belägg av ordet *egentligen* i inlärmaterialet faktiskt är satsadverbial, vilket skulle göra ordet det andra mest frekventa modala satsadverbial i hela inlärmaterialet.

Naturligtvis förekommer totalt 30 gånger i inlärmaterialet, och 87 % av träffarna har taggats som satsadverbial, medan det motsvarande talet i Talbanken är 91 %. De fyra träffarna som inte har kategoriserats som satsadverbial i inlärmaterialet är ändå i själva verket modala satsadverbial. Det totala antalet träffar på *naturligtvis* som satsadverbial skulle således vara 30, men ställningen som det femte mest frekventa modala satsadverbial i inlärmaterialet förblir oförändrad.

Satsadverbial	Absolut frekvens	Relativ frekvens	Rf (Bloggmix)
Kanske	178	1 133,8	827,4
Egentligen	65	414,0	235,0
Verkligen	59	375,8	587,5
Ju	49	312,1	2 210,7
Naturligtvis	30	191,1	47,3

Tabell 25. De fem mest frekventa modala satsadverbialen i hela inläromaterialet efter en manuell annotering, jämfört med de automatiskt genererade frekvenserna i Bloggmix-materialet.

Kontrollgranskningen visar att resultaten tydligen har påverkats av annoteringen. För det första skulle de fem mest frekventa modala satsadverbialen i hela inläromaterialet enligt den här granskningen vara i en annan ordning: *kanske* skulle fortfarande vara klart det mest vanligt förekommande modala satsadverbialen i materialet, men det skulle följas av *egentligen* och *verkligen* i respektive ordning, medan *ju* skulle komma först på fjärde plats. *Naturligtvis* skulle kvarstå på femte plats. Frekvensen av *naturligtvis* verkar nu sticka fram ännu mer i resultaten jämfört med den motsvarande siffran i Bloggmix-materialet, och frekvensen av *kanske* som tidigare var rätt lika med Bloggmix-materialet är nu tydligen högre. Även användningen av *egentligen* ger anledning för framtida forskning – medan det tidigare inte verkade finnas någon stor skillnad i användningen av det här ordet i inlärtexter och Bloggmix-materialet, visar siffrorna att *egentligen* faktiskt är nästan dubbelt så frekvent i inläromaterialet som i Bloggmix-materialet. Det är dock inte möjligt att utan förbehåll lita på att annoteringen i Bloggmix-materialet är absolut korrekt och inte innehåller några fel, trots att det handlar om förstaspråksmaterial som annoteringsprogrammet förmodligen hanterar bättre än andraspråksmaterial. Eftersom Bloggmix-materialet är så stort är det inte möjligt att göra en liknande granskning av annoteringskvalitet som gjordes med inläromaterialet, men ett stickprov visar att t.ex. ordet *egentligen* har totalt 221 496 träffar i Bloggmix-materialet, varav annoterade som satsadverbial är 144 690 belägg, vilket täcker bara 65 % av träffarna.⁴⁹ Med tanke på hur ordet i Talbanken alltid annoterades som satsadverbial, kan detta ge anledning till att förmoda att även i Bloggmix-materialet är annoteringen förmodligen inte fullständig lyckad.

⁴⁹ Sökning 24 (bilaga 1)

Kontrollgranskningen visade att annoteringen i inlärmaterialet verkar vara bristfällig, vilket inte minst kan bero på texttypen: okorrektheter med t.ex. rättstavningen och ordföljden, som är typiska för andraspråksmaterial, kan vara förvirrande för annoteringsprogram som ofta förtränas med förstaspråksdata (van Rooy 2015: 89). Exempel på felstavade ord som inte igenkänts av annoteringsprogrammet finns i exemplen 70 och 71 nedan. I exempel 70 har annoteringsprogrammet annoterat satsadverbialet *faktiskt* som substantiv och dess dependensrelation har markerats vara subjekt. Ordets grundform och lemgram är också tomma, vilket betyder att inte ens en sökning efter lemgrammet *faktisk* skulle ha visat ordet i resultaten. Det är dock inte alltid lika entydigt att avgöra vilket ord som informanten har åsyftat att använda: i exemplet 71 har informanten förmodligen tänkt på satsadverbialet *verkligen* men den ortografiska formen har blivit *verkling*. Det här ordet har också taggats som substantiv och kategoriserats som subjektspredikativ, och således en sökning efter satsadverbial visar inte heller det här belägget.

(70) Vi vet **faktist** inte vad kan hända i framtiden, men man får alltid göra en prognos.
(Student W20)

(71) Det var **verkling** en vänligt dag. (Student S82)

Det är svårt att säga hur stort antal modala satsadverbial som saknas i den här studiens resultat på grund av den missledande ortografiska formen. Det skulle dock överhuvudtaget vara svårt att dra gränsen mellan vilka felstavade ord som är tillräckligt nära det korrekta stavningen och kunde således räknas med i resultaten, och vilka som borde uteslutas eftersom man inte kunde vara absolut säker om vilket ord som informanten har åsyftat att använda. Då borde man dessutom diskutera om informanten kan överhuvudtaget sägas ha tillägnat sig ordet fullständigt om den skriftliga formen inte är korrekt – t.ex. Nation (2001: 28) hävdar att en förutsättning för att ett ord kan anses vara inlärd är att inläraren kan stava det korrekt.

Exemplet 72 visar ett fall där datorn kan ha blivit missledd av en felaktig ordföljd. Som konstaterades i kapitel 2.1.2, kan ordet *verkligen* inte stå i satsens fundament. Eftersom meningen i exempel 72 ändå inleds med *verkligen* kan annoteringsprogrammet ha tolkat att det inte kan fungera som modalt satsadverbial. *Verkligen* i det här fallet har annoterats som negerande satsadverbial, en kategori som enligt Teleman (1974: 236) innehåller ord som negerar satsinnehållet, t.ex. *inte*, *icke*, *aldrig* och *ej*.

(72) **Verkligen** dålig dagen för mig när jag förlorade min syster när min syster dog. (Student S18)

Det har blivit tydligt att resultaten påverkas av fel som kan spåras till den automatiska annoteringen, och bristfälligheter och felannoteringar gör det svårt att analysera resultat på ett pålitligt sätt. För att kunna dra exakta slutsatser om användningen av modala satsadverbial i L2-inlärartexter skulle det alltså vara nödvändigt att gå noggrant igenom resultaten och stryka de ord och uttryck som inte är satsadverbial, och därtill på något enhetligt sätt säkerställa att alla de ord och uttryck som är modala satsadverbial skulle bli medräknade i resultaten. Det är kanske nödvändigt att fråga om det alltså vore lättare att operera utifrån ett oannoterat material och göra annoteringen manuellt från första början.

I det här avsnittet har jag presenterat hur den automatiska annoteringen har påverkat resultaten i den här undersökningen. Det är klart att den automatiska annoteringen har haft många fördelar med tanke på just den här undersökningen: studiens genomförande har varit effektivt, enhetligt och genomskinligt. Jag har dessutom kunnat använda en större mängd data än vad som annars kunde ha varit möjligt. De här fördelarna bör ändå övervägas med de oundvikliga nackdelarna som nu har kommit fram.

5.6 Sammanfattning

Undersökningen visade att användningen av modala satsadverbial ökar jämnt i takt med att språkkunskaperna i övrigt förbättras – ju högre kunskapsnivå som studenten uppnår desto mer satsadverbial används i texterna och desto mer variation finns det i användningen av satsadverbial. I ljuset av frekvenserna används modala satsadverbial mest på nivå C2 (rf=6445,7) och minst på nivå A1 (rf=580,0). Det är ändå inte oproblemiskt att dra slutsatser utifrån de automatiskt genererade frekvenserna eftersom en närmare granskning på resultaten visade att det t.ex. på nivå A1 inte förekommer ett enda modalt satsadverbial, utan de ord som taggats som satsadverbial är felaktigt annoterade. C2 var däremot inte direkt jämförbar med andra nivåer eftersom materialet täckte bara få texter på den nivån. På grund av det låga antalet texter verkade det t.ex. bara vara två olika modala satsadverbial som systematiskt används i texterna på nivå C2, något som utan tvivel inte skulle stämma i ett större material.

Det enda modala satsadverbial som tycks användas i inlärmaterialet från och med nybörjarnivån är *kanske*, som dock inte alls används på nivå A1 och är relativt sällsynt också på nivå A2. *Kanske* visade sig vara det mest frekventa modala satsadverbial på alla kunskapsnivåer, vilket kunde delvis förklaras med ordets semantiska och pragmatiska egenskaper, och delvis med det stora antalet olika funktioner som ordet verkade kunna uppvisa i inlärtexter. Av andra modala satsadverbial förekommer *ju* mycket i inlärtexter från och med nivå B1. På nivåerna B1 och C1 är *ju* klart det andra mest frekventa modala satsadverbial i texterna, medan på nivå B2 föregås det av ordet *verkligen*. Den nedgående frekvensen av *ju* mellan nivåerna B1 och B2 kan möjligtvis förklaras med att ju högre kunskapsnivå som inläraren når, desto bättre behärskar hen det språkliga registret som bör användas i provtillfällen, vilket bidrar till att texterna blir mer formella och mindre talspråkliga. Att användningen av *ju* ökar igen på nivå C1 strider dock mot den här teorin, men det är möjligt att den igen ökande användningen beror på att *ju* på högre kunskapsnivåer uttrycker finare stilistiska nyanser. En kvalitativ analys av ordet *ju* visade flera olika funktioner för ordet i inlärmaterialet, varav några tydligen mer implicita än andra. En närmare kvalitativ undersökning på vilka av de här funktionerna som används på olika kunskapsnivåerna behövs dock för att kunna säga om det verkligen är så att avancerade andraspråksinlärare behärskar och utnyttjar de här finare stilistiska nyanserna mer än inlärare på lägre kunskapsnivåer.

I läroboksmaterialet är *kanske* och *ju* de två mest frekventa modala satsadverbialen på alla kunskapsnivåer, och deras frekvenser är betydligt mycket större jämfört med andra modala satsadverbial i läromaterialen från och med nivå A1 till och med nivå B1. På nivå B2 börjar användningen av *kanske* och *ju* jämnas ut och distributionen mellan de fem mest frekventa modala satsadverbialen blir mer balanserad. Generellt verkar användningen av modala satsadverbial i läromaterialen börja i större utsträckning på nivå A2, nå sin höjdpunkt på nivå B1, och sedan visa en nedgående trend på nivåerna B2 och C1.

Jämförelsen till oavhängiga förstaspråksdata, i det här fallet bloggtexter, visade att de fem mest frekventa modala satsadverbialen är delvis samma i bloggtexter och i inlärtexter. T.ex. är *ju* och *kanske* de vanligaste modala satsadverbialen i bloggtexter såsom de också är i inlärtexter, men medan *kanske* är det klart mest frekventa modala satsadverbial i inlärtexter, har ordet *ju* en särdeles hög frekvens jämfört med de andra modala satsadverbialen i bloggmaterialet, vilket kan vara ett för texttypen utmärkande drag.

Genom att jämföra resultaten i inlärmaterialet och i L1-materialen märktes det att användningen av ordet *naturligtvis* verkar vara flitigare i inlärtexter än i L1-materialen, medan användningen av *väl* och *gärna* var väldigt marginell i inlärtexter trots att de båda var relativt vanligt förekommande ord i L1-materialen. De här iakttagelserna kan tyda på en tendens hos andraspråksinlärare att å ena sidan överanvända ordet *naturligtvis* och å andra sidan att underanvända orden *gärna* och *väl*, men inga absoluta slutsatser om den möjliga över- eller underanvändningen av satsadverbial visade sig vara möjliga att dra på basis av det här materialet.

Till slut granskades annoteringens påverkan i resultaten med hjälp av kontrollsökningar som visade att annoteringen verkligen hade en stor påverkan på hur resultaten ser ut. Därför bör man förhålla sig kritiskt till de resultaten som kom fram i den här undersökningen, och de kan snarast ses som riktgivande.

6. DISKUSSION

I det här kapitlet diskuterar jag studiens resultat i förhållande till tidigare forskning samt tar ställning till studiens begränsningar. Till slut ger jag även förslag för framtida forskning.

6.1 Studiens resultat

I den här undersökningen har jag studerat andraspråksinlärnarnas användning av modala satsadverbial i svenska. Syftet var att utreda vilka modala satsadverbial som används på olika kunskapsnivåer och se hur användningen av modala satsadverbial utvecklas under språkinlärningsprocessen. Vidare jämfördes användningen av modala satsadverbial i inlärmaterialet med två typer av förstaspråksmaterial. Å ena sidan jämfördes L2-inlärttexter med lärobokstexter för att se vilka modala satsadverbial andraspråksinlärare överhuvudtaget exponeras för i andraspråksundervisningen. Å andra sidan utreddes användningen av modala satsadverbial i ett oavhängigt förstaspråksmaterial, bloggtexter, för att se hur användningen av modala satsadverbial i inlärnarspråket förhåller sig till den faktiska språkanvändningen hos modersmålstalare. Till slut granskade jag hur automatiskt annoterade korpusar lämpade sig för de här jämförelserna och evaluerade hur mycket felannoteringar sist och slutligen påverkade studiens resultat.

Som studiens första forskningsfråga ville jag utreda vilka modala satsadverbial som förekommer i inlärttexter på olika kunskapsnivåer. Resultaten visade att användningen av modala satsadverbial verkar börja på nivå A2 med ordet *kanske*, och det kan således sägas vara det första modala satsadverbial som andraspråksinlärare tillägnar sig. Resultaten stämmer således för den del överens med Hancock och Sanells studie (2009: 10) om adverb användning i L2-franska där *peut-être* (sv. kanske) visade sig förekomma i inlärnarspråket redan på nybörjarnivå. I mitt material används *kanske* dock flitigt även när språkkunskaperna förbättras – det är klart det mest frekventa modala satsadverbial på alla kunskapsnivåer. Ett annat modalt satsadverbial som är mycket frekvent på nästan alla kunskapsnivåer är *ju* som är det andra mest frekventa modala satsadverbial på

nivåerna B1 och C1 och det tredje mest frekventa på nivå B2. Den höga frekvensen av *ju* antar jag inte minst beror på dess höga frekvens i L1-svenskan, speciellt i talat språk. En grundläggande princip inom den bruksbaserade lingvistikens är att en individs språkbruk påverkas av input, dvs. ju fler gånger inläraren möter ett ord desto mer troligt det är att hen också själv börjar använda det (se också avsnitt 2.2.5). I och med att inläraren exponeras för ordet *ju* både i undervisningsmaterialen, såsom det kom fram i den här studien, och också utanför klassrummet, är det inte överraskande i sig att ordet *ju* används mycket också i inlärttexter. Vad som däremot inte var förväntat var hur frekvent ordet var också på högre kunskapsnivåer där texterna kunde antas vara mer formella och innehålla således färre för talspråk utmärkande drag, för att såsom även Bardel (2016: 79) påpekar, ju mer avancerad inläraren är, desto bättre borde hen känna till de olika semantiska och pragmatiska egenskaperna hos ord som kan bl.a. avgöra hur pass lämpligt ett ord är i en viss kontext. Att *ju* ändå är ett mycket vanligt modalt satsadverbial i inlärttexter på avancerad kunskapsnivå kan således möjligtvis bero på ordets finare stilistiska nyanser som inläraren kan utnyttja när inlärningsprocessen har kommit tillräckligt långt.

Förutom *kanske* och *ju* finns det lite variation i vilka andra modala satsadverbial som förekommer bland de vanligaste på olika kunskapsnivåer (se tabell 26). Det är viktigt att iaktta att i det här materialet är nivåerna A1 och C2 mycket mindre representerade än de andra nivåerna, och således är resultaten på de nivåerna inte direkt jämförbara med de andra nivåerna.

A1	A2	B1	B2	C1	C2
-	Kanske (419,8)	Kanske (710,7)	Kanske (1084,5)	Kanske (1028,1)	Kanske (3069,4)
-	-	Ju (375,3)	Verkligen (376,3)	Ju (523,8)	Tyvärr (613,9)
-	-	Naturligtvis (209,0)	Ju (243,5)	Egentligen (426,8)	-
-	-	Tyvärr (167,2)	Naturligtvis (177,1)	Enligt (388,0)	-
-	-	Verkligen (125,4)	Förstås/egentligen (132,8)	Verkligen (349,2)	-

Tabell 26. De fem mest frekventa modala satsadverbialen på olika kunskapsnivåer i inlärtmaterialet.

Den andra forskningsfrågan gällde möjliga över- eller underanvändningsmönster i användningen av modala satsadverbial hos andraspråksinlärare. I resultaten kom det fram att det finns ett ord som möjligen överanvänds av andraspråksinlärare, nämligen *naturligtvis*. Ordets frekvens i hela inlärmaterialet var 165,6, medan det motsvarande talet var 87,3 i läroboksmaterialet och 47,3 i bloggmaterial. Av de olika kunskapsnivåerna i inlärmaterialet användes ordet på nivå B1 (rf=209,0), på nivå B2 (rf=177,1) och på nivå C1 (rf=232,9). Frekvenserna är totalt sett inte höga men det är ändå fråga om ett mångfaldigt antal träffar av ordet i inlärmaterialet jämfört med L1-materialen, vilket absolut ger anledning till fortsatt forskning om ämnet.

Orden *gärna* och *väl* däremot dök upp mycket i bloggmaterial men var sällsynta i inlärtexter, trots att båda orden användes systematiskt även i lärobokstexter från och med nivån A2. Det här kunde eventuellt ses som ett tecken på att de här två orden underanvänds av andraspråksinlärare, men det är inte en entydig slutsats. Den låga frekvensen av *gärna* i inlärtexter antar jag kanske snarast berätta om ordets snäva användningsområde, medan den högre frekvensen av *väl* i bloggtexter möjligtvis beror på det faktumet att modersmålstalare helt enkelt känner till fler av ordets många funktioner än andraspråksinlärare. Den kvalitativa analysen visade nämligen att *väl* kan uppvisa flera olika typer av funktioner, trots att ordet i inlärmaterialet bara användes i en funktion: för att uttrycka förmodan. Att *väl* verkligen kan uppvisa ett antal olika funktioner har noterats också i tidigare forskning om ämnet, t.ex. av Aijmer (2015: 178, 192) som i sin kontrastiva analys mellan svenska, engelska och franska fann totalt 36 möjliga översättningar för *väl* i engelska texter och 34 i franska texter. Aijmer (2015: 198) framhäver ordets polysemiska karaktär, vilket jag antar kan vara förvirrande för andraspråksinlärare. Granger och Tyson (1996: 24–25) poängterar i sin studie om konnektorbruket i L2-engelska att de syntaktiska och semantiska skillnaderna mellan olika konnektorer kan vara små och svåra att förstå för andraspråksinlärare, och även mera avancerade studenter kan bli missledda av dem. Den här iakttagelsen påstår jag att gäller också för modala satsadverbial.

Något absolut svar på frågan om möjliga över- eller underanvändningstendenser av satsadverbial hos andraspråksinlärare var alltså inte möjligt att ge på basis av det här materialet, eftersom den avvikande användningen av orden kan helt enkelt bero på olikheter i texttypen mellan bloggtexter och inlärtexter. Användningen av orden *naturligtvis*, *gärna* och *väl* kunde ändå vara värt att studera vidare, inte minst genom att

jämföra L2-inlärartexter med exakt motsvarande data av modersmålstalare för att se om den flitiga användningen av *naturligtvis* och den ringa användningen av *gärna* och *väl* verkligen är ett för inlärarspråket utmärkande drag, eller om det helt enkelt bara är mer typiskt för L2-inlärartexter som texttyp än det är för bloggtexter.

Den tredje och sista forskningsfrågan gällde valet av att använda automatiskt annoterade korpusar som material i den här typen av undersökning. Användningen av annoterade och oannoterade korpusar är en vattendelare mellan olika korpuslingvistiska skolor, och motståndare till annoterade korpusar framhåller ofta de felaktiga eller oexakta annoteringar som kan förvrida resultat, speciellt om annoteringen är automatiskt gjord. De som förespråkar användningen av annoterade korpusar framhäver däremot vikten av genomskinliga processer så att de möjliga felen kan spåras till materialet. Dessutom är det naturligtvis tidsbesparande att använda annoterade korpusar, vilket i dagens värld är en oneklig fördel. Jag ville bidra till den diskussionen som råder kring de olika typer av korpusarna genom att granska hur valet av att använda automatiskt annoterade korpusar påverkade just den här studiens resultat.

Efter att ha utrett hur korrekt annoteringen av den här studiens inlärmaterial är, märkte jag att resultaten som kom fram om användningen av modala satsadverbial kan snarast ses som riktgivande, eftersom annoteringen inte var fullständig lyckad. Till de av korpusverket producerade resultaten hade bl.a. medräknats ord som faktiskt inte är modala satsadverbial, i några fall inte ens satsadverbial, ibland till och med inte ens adverbial. Däremot saknades några modala satsadverbial i resultaten eftersom de hade kategoriserats fel. För att kunna rätta ut de här snedvridningarna borde jag ha gått igenom alla texterna manuellt, och då är det kanske befogat att fråga om det överhuvudtaget skulle ha varit lättare att jobba utifrån en rå korpus. Svaret är delvis ja. Om syftet är att forska i användningen av modala satsadverbial, är det obestriddligen ett seriöst problem att resultaten för det första inte omfattar alla modala satsadverbial i materialet, och dessutom omfattar sådana ord som inte ens borde ingå i resultaten.

Det är alltså inte möjligt att säga att de resultaten som kom fram i undersökningen skulle vara exakta, kanske inte ens nära exakta. Det är ändå värt att granska de ovannämnda nackdelarna i förhållande till de fördelarna som finns med användningen av automatiskt annoterade korpusar.

För det första, skulle det naturligtvis ha krävt mer tid och resurser att jobba utifrån ett oannoterat material. Det är inte heller självklart att en manuell annotering av materialet skulle ha lett till absolut felfria resultat, utan det kunde ha också innehållit fel och bristfälligheter, inte minst på grund av materialets storlek. Om den färdiggjorda annoteringen inte var tillgänglig skulle jag dock inte ens ha kunnat använda ett så här stort material, utan någon slags selektion av data borde ha gjorts, vilket som konstaterades i kapitel 4.3.1, inte är önskvärt inom korpuslingvistik. En undersökning som skulle ha baserat sig på ett mindre material kunde ha gett mer exakta resultat, men hur representativa skulle de resultaten då ha varit? Även processernas genomsnittlighet är en obestridd fördel med det här materialvalet. Resultaten är lätta att reproducera – i princip kan vem som helst som har tillgång till korpusarna göra samma sökningar och få samma resultat. Också de felaktigheterna som finns i resultaten är lätta att spåra till materialet, något som kunde vara svårt om den annoteringen som utnyttjades inte var tillgänglig för alla. Arbetandet med annoterade korpusar har dessutom varit lätt och effektivt, något som inte bör underskattas.

Sammanfattningsvis är det möjligt att konstatera att det absolut finns problematik med den automatiska annoteringen och att resultaten av den här undersökningen således ska tolkas med tanke på det. Det är självklart att teknologin förbättras hela tiden och att även annoteringsprogrammen kommer i framtiden att kunna producera mer tillförlitliga resultat. Innan den utvecklingen har kommit tillräckligt långt, skulle det ideala sättet att närma sig sådana forskningsfrågor som beskrivs i den här avhandlingen kanske vara att använda åtminstone delvis manuellt annoterade korpusar, vilket dock förmodligen skulle ställa upp striktare gränser för hur stora mängder data som kunde analyseras. Andraspråksmaterial som innehåller fel och avvikelser från L1-varietetet är fortfarande en utmaning för annoteringsprogram, och det är självklart att det goda med den automatiska annoteringen inte på riktigt kommer till sin rätt med just andraspråksmaterialet. Jag vågar ändå påstå att om man tittar på helhetsbilden, överträffar den automatiska annoteringens fördelar dess nackdelar också i den här undersökningen, och trots att några exakta slutsatser inte kan fastställas, ger studien en bra grund för framtida forskning.

6.2 Studiens begränsningar

Det är nödvändigt att lägga fram några begränsningar som gäller studiens material och metod och som ska tas i beaktande när resultaten tolkas och evalueras.

För det första är det viktigt att nämna materialets ofullständighet. En tillförlitlig jämförelse av språkanvändningen mellan olika kunskapsnivåer är inte möjligt med ett alltför ojämnt antal texter från olika nivåer. I min undersökning är inläratexterna på nivåerna A1 och C2 mycket mindre representerade än de andra nivåerna. Antalet texter på nivå A1 är 55 och antalet tokens 5172. Medellängden för texterna är alltså ca 94 tokens. Nivå C2 representeras däremot av bara sju enskilda texter, men eftersom det totala antalet tokens på nivå C2 är 3258, blir medellängden för en text ca 465 tokens. Varken ett stort antal enskilda texter eller en relativt hög medellängd hos dem garanterar nödvändigtvis att materialet skulle vara tillräckligt representativt. Jämfört med t.ex. inlärmaterialet på nivå C1 som består av 51 549 tokens och 96 enskilda texter vars medellängd är således ca 537 tokens, blir det tydligt att varken A1 eller C2 är direkt jämförbar med resten av nivåerna.

På grund av den ringa mängden material på nivåerna A1 och C2 är det överhuvudtaget svårt att säga någonting om användningen av modala satsadverbial på de här nivåerna. Som det kom fram i avsnitt 5.2.1 förekommer det, åtminstone om man lutar på den automatiska annoteringen, inga modala satsadverbial i inläratexterna på nivå A1. Även ett snabbt stickprov verkar bekräfta den här hypotesen: en enkel sökning efter de fem mest frekventa modala satsadverbialen i hela inlärmaterialet (*kanske, ju, verkligen, egentligen och naturligtvis*), ger inte några träffar i inlärmaterialet på nivå A1.⁵⁰ Det skulle ändå vara farligt att dra den slutsatsen att modala satsadverbial inte alls skulle förekomma i inlärspråket på nivå A1 trots att det visade sig vara så i det här materialet: ett enda exempel mot den här hypotesen skulle räcka för att vederlägga den, såsom exemplifieras även av McEnery och Hardie:

Given the hypothesis that all planets have moons, if we have data available from a thousand planets, the fact that 999 of them have moons is not as important – from the point of view of defending the hypothesis – as the fact that one planet has no moons at all. Likewise, if the hypothesis we are looking is that some particular

⁵⁰ Sökning 25 (bilaga 1)

linguistic form never occurs, then the only part of the corpus that is really relevant is the part where that linguistic form *does* occur, thus falsifying the hypothesis. (McEnery & Hardie, 2012: 17.)

På nivå C2 finns det enligt den automatiska annoteringen bara två modala satsadverbial som mer eller mindre systematiskt användes i texterna, nämligen *kanske* och *tyvärr*. En kontrollsökning i den delen av inlärmaterialet som representerar nivå C2 delvis bekräftar hypotesen: den ger inga träffar för *egentligen* och bara en träff för *ju* och *naturligtvis* vardera.⁵¹ En sökning efter *verkligen* ger däremot 3 träffar, varav två i en och samma text. Att det finns ett tredje belägg som felaktigt annoterats som *annat adverbial* och som således saknas i de automatiskt genererade resultaten betyder att även *verkligen* används av fler än en informant på nivå C2. Antalet förekomster av modala satsadverbial är dock i varje fall tydligen för litet på nivå C2 för att kunna på ett pålitligt sätt berätta om användningen av modala satsadverbial på den nivån.

En begränsning som redan diskuterades i samband med materialbeskrivningen gäller materialens jämförbarhet. Att inlärttexter jämförs med bloggtexter är inte idealt, men med hänsyn till de tillgängliga resurserna var det ändå ett motiverat val. Tanken var att eftersom exakt motsvarande förstaspråksdata inte var tillgängliga, skulle jämförelsematerialet åtminstone representera det naturliga L1-språkbruket, som inte skulle vara för vårdat, professionellt eller genomtänkt. Vidare ville jag välja jämförelsematerialet så att det inte skulle representera en homogen texttyp som skulle präglas för mycket av just för den texttypen utmärkande drag. Dessutom kunde jag genom att välja ett jämförelsematerial från just Språkbankens korpusar göra alla sökningar i de tre delmaterialen med hjälp av korpusverktyget Korp, och resultaten blev således enhetliga och lätta att reproducera. Eftersom alla de tre delmaterialen är automatiskt annoterade enligt samma kriterier, är annoteringen även konsekvent, vilket som påpekades i avsnitt 4.1, också är ytterst viktigt. På basis av de här kriterierna valdes Bloggmix-korpusarna som jämförelsematerial för inlärttexter, trots att språkanvändningen i de här två texttyperna kunde förväntas vara annorlunda, och således är det inte lätt att konstatera vilka skillnader som beror på olikheter mellan inlärtarspråket

⁵¹ Sökning 26 (bilaga 1)

och målspråket, och vilka som helt enkelt kan förklaras med olikheter mellan inlärttexter och bloggtexter som texttyp.

Som diskuterades tidigare kan studiens resultat inte tolkas vara exakta, eftersom den automatiska annoteringen visade sig påverka relativt mycket på resultaten. Detta exemplifierades med hjälp av kontrollsökningar där de riktiga antalen av de fem mest frekventa modala satsadverbialen i hela inlärtmaterialet räknades. Det faktumet att det visade sig faktiskt finnas fler träffar på de här fem satsadverbialen än vad som ingår i resultaten förvrider resultat också när det gäller satsadverbialanvändningen på enskilda nivåer, såsom framgår av tabell 27 som visar de fem mest frekventa orden som har taggats som satsadverbial i inlärtmaterialet på nivå C1, samt frekvenserna när de här fem orden söktes separat i materialet, alltså utan den begränsningen att ordet måste ha taggats som satsadverbial. Resultaten är beaktansvärda: om alla de belägg som inte har taggats som satsadverbial var faktiskt satsadverbial, skulle t.ex. ordet *enligt* vara det mest frekventa modala satsadverbialet i inlärttexter på nivå C1, vilket skulle bl.a. vederlägga hypotesen att *kanske* är det mest frekventa modala satsadverbialet på alla kunskapsnivåer. Det är dock möjligt att inte alla de träffar som har taggats som någonting annat än satsadverbial faktiskt skulle vara satsadverbial, men ändå är det värt att beakta den möjligheten.

Satsadverbial	Antal träffar (annoterad som satsadverbial)	Antal träffar (alla träffar)	% av träffar som annoterats som satsadverbial
Kanske	53 (rf=1028,1)	67 (rf=1 299,7)	79 %
Ju	27 (rf=523,8)	28 (rf= 543,2) ⁵²	96 %
Egentligen	22 (rf=426,8)	51 (rf= 989,3)	43 %
Enligt	20 (rf=388,0)	88 (rf=1 707,1)	22 %
Verkligen	18 (rf=349,2)	25 (rf=485,0)	72 %

Tabell 27. Frekvenserna av de fem mest frekventa modala satsadverbialen i inlärttexter på nivå C1, annoterade som satsadverbial versus annoterade med vilken som helst dependensrelation.⁵³

Den här undersökningen beaktar inte i större utsträckning den individuella variationen i användningen av modala satsadverbial. Trots att jag under studiens gång har varit noga

⁵² Förekomster av konjunktionen *ju-desto* har räknats bort.

⁵³ Sökning 27 (bilaga 1)

med att se till att adverbial som tas med i resultaten används av fler än en informant, är det möjligt att några satsadverbial som förefaller vara frekventa i inlärmaterialet i själva verket bara används flera gånger av några enskilda informanter. I en mer omfattande forskning borde också den här aspekten tas i beaktande för att säkerställa att resultaten i verkligheten är representativa på gruppnivå. Det var inte heller möjligt att utvärdera hur transfer påverkar resultaten, m.a.o. hur användningen av modala satsadverbial kan reflektera ett visst modersmåls påverkan. T.ex. i Granger och Tysons studie (1996: 22–23) kom det fram att det engelska uttrycket *on the contrary* överanvändes av franska andraspråksinlärare på grund av att det felaktigt användes i samma kontexter som det franska uttrycket *au contraire*. Sådana här samband är i princip helt möjliga också i mitt material, och det är självklart att om över- eller underanvändningen av något satsadverbial beror på ett visst modersmåls påverkan, borde det inte tolkas som ett för hela inlärarspråket kännetecknande drag. Det är också möjligt att vissa modala satsadverbial är av materialspecifika skäl mer representerade än andra, såsom märktes t.ex. i användningen av ordet *enligt*.

Till slut vill jag ännu konstatera att även det här materialet är ett finit sampel av en infinit enhet, dvs. språk. Materialet ger därför bara en begränsad syn på användningen av satsadverbial, men ju mer forskning om ämnet bedrivs desto mer kännedom kan vi få om det, och desto mer pålitliga slutsatser är det möjligt att dra.

6.3 Slutsatser och förslag för framtida forskning

Den här undersökningen ger en bra grund för framtida forskning om ämnet modala satsadverbial i inlärarspråket, trots att exakta slutsatser på basis av det här materialet inte kunde fastställas. Forskning om satsadverbial kan ge viktig information för de som jobbar inom språkutbildningen – inte minst lärare och läroboksförfattare – och det kan hjälpa dem i deras arbete för att kunna bättre stödja andraspråksinlärare på vägen mot deras fullständiga språkliga potential. Behärsksningen av modala satsadverbial kan behövas för att överhuvudtaget klara sig i samtal med modersmålstalare, men en skicklig användning av modala satsadverbial kan även liva upp språkbruket och ge en uppfattning om en högutvecklad språkförmåga, vilket kan vara en stor nytta t.ex. på arbetsmarknaden för människor med andraspråksbakgrund. Jag tror också att ju mer övertygad

andraspråksinläraren är om sin egen förmåga att förstå och också själv uttrycka även de finare nyanserna i ett språk, desto mer troligt är det att hen vågar använda språket också utanför klassrummet. Känslan av att kunna förstå och själv bli förstådd kan för sin del vara av avgörande betydelse t.ex. i integration till ett främmande land.

Den här avhandlingen beskrev användningen av modala satsadverbial främst ur ett kvantitativt perspektiv, men även den kvalitativa delen behandlades ytligt. I framtiden kunde den kvalitativa analysen fördjupas t.ex. genom att jämföra L2-inlärartexter med motsvarande texter av modersmålstalare för att se vilka alla funktioner som andraspråksinlärare uttrycker med modala satsadverbial, och vilka funktioner som förekommer i L1-texter men inte i L2-texter. Den jämförelsen kunde eventuellt ge kunskap om vilka funktioner som borde tas upp mer i andraspråksundervisningen för att göra andraspråksinlärare medvetna om satsadverbialens fullständiga betydelsepotential. Att studera närmare *hur* andraspråksinlärare använder modala satsadverbial i olika språkliga kontexter skulle även vara viktigt med tanke på det faktumet att den korrekta användningen av modala satsadverbial inte bara handlar om grammatisk korrekthet, utan det är minst lika viktigt att känna till de stilistiska nyanserna som modala satsadverbial realiserar i olika kontexter. Vikten av att fästa uppmärksamhet vid de olika språkliga registren i andraspråksundervisningen framhävs även av Jiménez Catalán och Ojeda Alba (2010: 92).

Den här studien ger även anledning till nya kvantitativa studier om användningen av modala satsadverbial i L2-svenskan. Inte minst skulle det vara viktigt att jämföra den här studiens resultat med ett mer jämförelsebart L1-material för att utreda om t.ex. ordet *naturligtvis* verkligen överanvänds av andraspråksinlärare, och om orden *väl* och *gärna* underanvänds av dem, såsom den här studiens resultat låter förstå. Det vore även intressant att gå närmare in på ordinlärningsprocessen och t.ex. ta reda på i vilket skede av språkinlärningsprocessen som enskilda satsadverbial tillägnas i det receptiva ordförrådet, och hur länge det tar tills de etableras också i det produktiva ordförrådet. Eftersom den här avhandlingen har fokuserat på skriftligt material, skulle det dessutom behövas forskning om användningen av modala satsadverbial i talad andraspråkssvenskan.

Jag hoppas att den här avhandlingen har även bidragit till diskussionen kring användningen av annoterade korpusar genom att på ett konkret sätt ha visat vilka för- och

nackdelar som finns med automatiskt annoterade korpusar i den här typen av undersökning. Vidare hoppas jag att i och med att ha visat vad som redan är möjligt att göra med automatiskt annoterade korpusar och vad som fortfarande skapar utmaningar för den automatiska annoteringen, har den här avhandlingen kunnat ge viktig information för de som jobbar i utvecklingen av korpuslingvistiska verktyg, så att det i närmaste framtiden blir möjligt att få ännu mer pålitliga resultat med den automatiska annoteringen, vilket för sin del kommer att öppna helt nya dörrar för andraspråksforskningen.

LITTERATUR

- Abrahamsson, N. 2009. *Andraspråksinläring*. Lund: Studentlitteratur.
- af Hällström-Reijonen, C. 2014. Regionala finlandismer i samtida finlandssvenskt tidningsspråk. I: Kosunen, R., Lepistö, K. & Rossi, P. 2014. *Svenskan i Finland: 14*. Oulu: Oulu Universitet.
- Aijmer, K. 2016. The Swedish modal particle *nog*: A contrastive analysis. *Nordic Journal of English Studies: NJES*, 15(3), 149–170. <https://doi.org/10.35360/njes.379>
- Aijmer, K. 2015. The Swedish modal particle *väl* in a contrastive perspective. *Nordic Journal of English Studies: NJES*, 14(1), 174–200. <http://doi.org/10.35360/njes.344>
- Aijmer, K. 1996. Swedish modal particles in a contrastive perspective. *Language Sciences (Oxford)*, 18(1), 393–427. [https://doi.org/10.1016/0388-0001\(96\)00027-7](https://doi.org/10.1016/0388-0001(96)00027-7)
- Andréasson, M. 2007. *Satsadverbial, ledföljd och informationsdynamik i svenskan*. Göteborg: Institutionen för svenska språket, Göteborgs universitet.
- Bardel, C. 2016. The lexicon of advanced L2 learners. I: Hyltenstam, K. 2016. *Advanced proficiency and exceptional ability in second languages*. Berlin: De Gruyter Mouton.
- Biber, D., Conrad, S. & Reppen, R. 1998. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press.
- Bolander, M. 1987. "Man kan studera inte så mycke": Om placering av negation och adverb i vuxna invandrades svenska. Stockholm: Stockholms universitet.
- Borin, L., Forsberg, M. & Roxendal, J. 2012. Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012*. 474–478. Istanbul: ELRA.
- Brake, D. R. 2009. *As if nobody's reading? The imagined audience and socio-technical biases in personal blogging practice in the UK*. London: London School of Economics.
- Enström, I. 2010. *Ordens värld: svenska ord – struktur och inläring*. 1. uppl. Uppsala: Hallgren & Fallgren.
- Europarådet. 2020. *Common European framework of reference for languages: learning, teaching, assessment – Companion volume*. Europarådet.
- Figueras, N. 2012. The impact of the CEFR. *ELT journal vol. 66 (4)*, 477–485. <https://doi.org/10.1093/elt/ccs037>
- Forsberg, U. 1988. *Modala satsadverbial i finlandssvenskt och rikssvenskt tidningsspråk*. Helsingfors: Helsingfors universitet.
- Gilquin, G., Granger, S. & Meunier, F. 2015. Introduction: learner corpus research – past, present and future. I: Gilquin, G., Granger, S. & Meunier, F. 2015. *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.
- Granger, S. 2015. Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7–24. <https://doi.org/10.1075/ijlcr.1.1.01gra>

- Granger, S. 2002. A Bird's-eye view of learner corpus research. I: Granger, S., Hung, J., Petch-Tyson, S. 2002. *Computer learner corpora, second language acquisition, and foreign language teaching*. Amsterdam; Philadelphia: John Benjamins Pub.
- Granger, S. & Tyson, S. 1996. Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17–27. <https://doi.org/10.1111/j.1467-971X.1996.tb00089.x>
- Hammarberg, B. 2004. Teoretiska ramar för andraspråksforskning. I: Hyltenstam, K. & Lindberg, I. 2004. *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur.
- Hancock, V. & Sanell, A. 2009. The acquisition of four adverbs in a learner corpus of L2 French. *Discours*, 5(5). <https://doi.org/10.4000/discours.7368>
- Harsch, C. 2018. How suitable is the CEFR for setting university entrance standards? *Language Assessment Quarterly*, 15(1), 102–108. <https://doi.org/10.1080/15434303.2017.1420793>
- Helsingfors stad. Kotoutuminen Suomeen. Hämtad 15.3.2021. <https://www.infofinland.fi/fi/elama-suomessa/asettuminen-suomeen/kotoutuminen-suomeen>
- Himberg, P. 2013. Kun pakollisesta kouluruotsista tuli ongelma. Hämtad 4.9.2021. <https://yle.fi/aihe/artikkeli/2013/09/26/kun-pakollisesta-kouluruotsista-tuli-ongelma>
- Hultman, T. G. 2003. *Svenska Akademiens språklära*. Stockholm: Svenska Akademien.
- Hyltenstam, K. 1978. *Progress in immigrant Swedish syntax: A variability analysis*. Lund: Institutionen för lingvistik, Lunds universitet.
- Hyltenstam, K. & Lindberg, I. 2004. *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur.
- Håkansson, G. 1989. The acquisition of negative placement in Swedish. *Studia Linguistica*, 43(1), 47–58. <https://doi.org/10.1111/j.1467-9582.1989.tb00794.x>
- Ivaska, I. 2014. Mahdollisuuden ilmaiseminen S1-suomea ja edistynyttä S2-suomea erottavana piirteenä. *Lähivertailuja* 24, 47–80. <https://doi.org/10.5128/LV24.02>
- Jiménez Catalán, R. M. & Ojeda Alba, J. 2010. Connectors in EFL learners' essays and in course books. I: Calzada Pérez, M., Moreno Jaén, M. & Serrano Valverde, F. 2010. *Exploring new paths in language pedagogy: lexis and corpus-based language teaching*. London: Equinox Pub.
- Josephson, O. 2005. *Ju: ifrågasätta självklarheter om svenskan, engelskan och alla andra språk i Sverige*. 2. uppl. Stockholm: Norstedts akademiska förlag.
- Kaye, B. L. 2010. Going to the blogs: Toward the development of a uses and gratifications measurement scale for blogs. *Atlantic Journal of Communication*, 18(4), 194–210. <https://doi.org/10.1080/15456870.2010.505904>
- Lagerholm, P. 2010. *Språkvetenskapliga uppsatser*. 2. uppl. Lund: Studentlitteratur.
- Lagerholm, P. 2008. *Stilistik*. Lund: Studentlitteratur.
- Lindström, J. 2008. *Tur och ordning*. Stockholm: Norstedts akademiska förlag.

- Lyngfelt, B. & Prentice, J. 2020. *Ett bruksbaserat perspektiv på språkinläring*. Göteborg: Institutionen för svenska språket, Göteborgs universitet.
- Lövheim, M. 2011. Personal and popular. The case of young Swedish female top-bloggers. *Nordicom Review*, 32(1), 3–16. <https://doi.org/10.1515/nor-2017-0101>
- McEnery, T. & Hardie, A. 2012. *Corpus linguistics: method, theory and practice*. Cambridge; New York: Cambridge University Press.
- Migrationsinfo.se. Hur många i Sverige är födda i ett annat land? Hämtad 15.7.2021. <https://www.migrationsinfo.se/fragor-och-svar/hur-manga-utrikes-fodda-sverige/>
- Muittari, V. 1988. *Om modus och modalitet: en finsk-svensk kontrastiv studie över modalitet i konditionala, koncessiva och komparativa satser och satsfogningar*. Jyväskylä: Jyväskylä universitet.
- Nation, I. S. P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nilsson, J. 2005. *Adverb i interaktion*. Göteborg: Institutionen för svenska språket, Göteborgs universitet.
- Norrby, C. & Håkansson, G. 2007. *Språkinläring och språkanvändning. Svenska som andraspråk i och utanför Sverige*. Lund: Studentlitteratur.
- Pérez-Paredes, P. 2010. The death of the adverb revisited: Attested uses of adverbs in native and non-native comparable corpora of spoken English. I: Calzada Pérez, M., Moreno Jaén, M. & Serrano Valverde, F. 2010. *Exploring new paths in language pedagogy: lexis and corpus-based language teaching*. London: Equinox Pub.
- Petersson, D. 2008. *Inte, nog och visst i mittfält och fundament*. Lund: Lunds universitet.
- Saeed, J. I. 2016. *Semantics*. 4. uppl. Chichester: Wiley Blackwell.
- Shimada, K. 2014. Contrastive interlanguage analysis of discourse markers used by nonnative and native English speakers. *JALT Journal* 36(1), 47–67. <https://doi.org/10.37546/JALTJJ36.1-3>
- Sigurd, B. & Håkansson, G. 2007. *Språk, språkinläring och språkforskning*. 2. uppl. Lund: Studentlitteratur.
- Skolverket. 2009. *Gemensam europeisk referensram för språk: lärande, undervisning och bedömning*. Stockholm: Skolverket.
- Språkbanken. Korps användarhandledning. Hämtad 21.7.2021. <https://spraakbanken.gu.se/en/tools/korp/usermanual>
- Språkläraryrket i Finland. *Perustietoa kielivalinnoista*. Hämtad 15.3.2021. https://www.sukol.fi/liitto/tilastot/perustietoa_kielivalinnoista
- Statistikcentralen. 2020. *Finland i siffror 2020*. https://www.stat.fi/tup/julkaisut/tiedostot/julkaisuluettelo/yyti_fis_202000_2020_232_13_net
- Svenska institutet. 2020. Svenska språket – en allmän beskrivning. Hämtad 15.3.2021. <https://svenskaspraket.si.se/svenska>
- Svensk ordbok (SO). 2021. Tillgänglig via <https://www.svenska.se>

- Svensson, G. 2009. *Diskurspartiklar hos ungdomar i mångspråkiga miljöer i Malmö*. Lund: Lunds universitet.
- Tandefelt, M. 2005. Ju. *Språkbruk 2/2005*.
- Taylor, C. 2008. What is corpus linguistics? What the data says. *ICAME Journal 32*, 179–200.
- Teleman, U., Andersson, E., Christensen, L. & Hellberg, S. 1999a. *Svenska Akademiens grammatik 1 – Inledning och register*. Stockholm: Svenska Akademien.
- Teleman, U., Andersson, E., Christensen, L. & Hellberg, S. 1999b. *Svenska Akademiens grammatik 4 – Satser och meningar*. Stockholm: Svenska Akademien.
- Teleman, U. 1974. *Manual för grammatisk beskrivning av talad och skriven svenska*. Lund: Studentlitteratur.
- Utbildningsstyrelsen. Undervisning i svenska som andraspråk. Hämtad 22.2.2021. <https://www.oph.fi/sv/utbildning-och-examina/undervisning-i-svenska-som-andrasprak>
- van Rooy, B. 2015. Annotating learner corpora. I: Gilquin, G., Granger, S. & Meunier, F. 2015. *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.
- Viberg, Å. 2004. Lexikal utveckling i ett andraspråk. I: Hyltenstam, K. & Lindberg, I. 2004. *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur.
- Volodina, E., Pilán, I., Enström, I., Llozhi, L., Lundkvist, P., Sundberg, G., Sandell, M. 2016. SweLL on the rise: Swedish learner language corpus for European reference level studies. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. 206–212. Portorož: ELRA. <https://aclanthology.org/L16-1031>
- Volodina, E., Pilán, I., Rødven Eide, S., Heidarsson, H. 2014. You get what you annotate: a pedagogically annotated corpus of coursebooks for Swedish as a second language. *Proceedings of the third workshop on NLP for computer-assisted language learning*. 128–144. Uppsala: LiU Electronic Press. <https://aclanthology.org/W14-3510>
- von Stutterheim, C. 1993. Modality: form and function in discourse. i: Dittmar, N. & Reich, A. 1993. *Modality in language acquisition=Modalité et acquisition des langues*. Berlin; New York: Gruyter.
- Wray, A. & Bloomer, A. 2012. *Projects in linguistics and language studies: a practical guide to researching language*. 3. uppl. London: Hodder Education.

Korpusar

Alla korpusar tillgängliga via Korp <https://spraakbanken.gu.se/korp>

Bloggmix 1998

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix1998>

Bloggmix 1999

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix1999>

Bloggmix 2000

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2000>

Bloggmix 2001

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2001>

Bloggmix 2002

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2002>

Bloggmix 2003

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2003>

Bloggmix 2004

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2004>

Bloggmix 2005

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2005>

Bloggmix 2006

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2006>

Bloggmix 2007

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2007>

Bloggmix 2008

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2008>

Bloggmix 2009

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2009>

Bloggmix 2010

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2010>

Bloggmix 2011

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2011>

Bloggmix 2012

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2012>

Bloggmix 2013

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2013>

Bloggmix 2014

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2014>

Bloggmix 2015

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2015>

Bloggmix 2016

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2016>

Bloggmix 2017

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmix2017>

Bloggmix okänt datum

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=bloggmixodat>

COCTAILL

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=coctail>

PAROLE

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=parole>

SpIn

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=spin-source>

Sw1203-uppsatser version 2

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=sw1203v2>

Talbanken

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=talbanken>

TISUS-texter v2

<https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=tisusv2>

BILAGA 1. Sökningarna i Korp

Sökning 1 – De mest frekventa modala satsadverbialen i inlärrarmaterialet

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex&search=cqp&result_tab=2

Sökning 2 – Kanske, ju, verkligen, egentligen och naturligtvis i inlärrarmaterialet (utan dependensrelation)

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex&search=cqp&result_tab=2

Sökning 3 – Användningen av modala satsadverbial per kunskapsnivå (inlärrarmaterialet)

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=text_grade&result_tab=2&search=cqp

Sökning 4 – De mest frekventa modala satsadverbialen per kunskapsnivå i inlärrarmaterialet (med nivåerna A1, A2, B1, B2, C1 och C2)

Sammanställt på lemgram:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20_text_grade%20%3D%20%22A1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

Sammanställt på ord:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20_text_grade%20%3D%20%22A1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&search=cqp

Sökning 5 – Inte i inlärrartexterna (med nivåerna A1, A2)

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%20%26%20_text_grade%20%3D%20%22A1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=word,text_grade&result_tab=2&search=cqp

Sökning 6 – De mest frekventa modala satsadverbialen per kunskapsnivå (COCTAILL) (med nivåerna A1, A2, B1, B2, C1)

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20_lesson_level%20%3D%20%22A1%22%5D&corpus=coctail&stats_reduce=lex&result_tab=2&search_tab=1&search=cqp

Sökning 7 – Inte på nivå A1 i COCTAILL

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%20%26%20_lesson_level%20%3D%20%22A1%22%5D&corpus=coctail&search_tab=1&search=cqp&stats_reduce=lex

Sökning 8 – Frekvensen av satsadverbial på olika kunskapsnivåer i COCTAILL

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=coctail&search_tab=1&result_tab=2&stats_reduce=lesson_level&search=cqp

Sökning 9 – De mest frekventa modala satsadverbialen i Bloggmix-materialet

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%5D&result_tab=2&search_tab=1&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmix2017,bloggmixodat&stats_reduce=lex&search=cqp

Sökning 10 – Modala satsadverbial i studiens tre delmaterial

Inlärrarmaterialet:

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%5D&search_tab=1&corpus=spin-source,sw1203v2,tisusv2&stats_reduce=lex&result_tab=2&search=cqp

COCTAILL:

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%5D&search_tab=1&corpus=coctail&result_tab=2&stats_reduce=lex&search=cqp

Bloggmix:

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%5D&search_tab=1&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmix2017,bloggmixodat&result_tab=2&stats_reduce=lex&search=cqp

Sökning 11 – Modala satsadverbial i inlärmaterialet vs bloggmaterial

Inlärmaterialet:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20_text_grade%20%3D%20%22A1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

Bloggmaterial:

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%5D&search_tab=1&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmix2017,bloggmixodat&result_tab=2&stats_reduce=lex&search=cqp

Sökning 12 – Tyvärr i inlärmaterialet

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22tyv%C3%A4rr%5C.%5C.ab%5C.1%22%5D&search_tab=1&corpus=spin-source,sw1203v2,tisusv2&stats_reduce=lex,text_grade&result_tab=2&search=cqp

Sökning 13 – Inte i inlärmaterialet

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%5D&search_tab=1&corpus=spin-source,sw1203v2,tisusv2&stats_reduce=lex&search=cqp&result_tab=2

Sökning 14 – Tyvärr i COCTAILL

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22tyv%C3%A4rr%5C.%5C.ab%5C.1%22%5D&search_tab=1&corpus=coctail&result_tab=2&stats_reduce=lex&search=cqp

Sökning 15 – Tyvärr i Bloggmix

https://spraakbanken.gu.se/korp/#?%2F%3Fsearch_tab=1&cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22tyv%C3%A4rr%5C.%5C.ab%5C.1%22%5D&search_tab=1&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmixodat&stats_reduce=lex&search=cqp&result_tab=2

Sökning 16 – Enskilda informanter och satsadverbial

Orden kanske, ju, verkligen, egentligen, naturligtvis:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_student_id&result_tab=2&search=cqp

Ordet inte:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_student_id&result_tab=2&search=cqp

Sökning 17 – Enskilda inlärartexter och satsadverbial

Kanske, ju, verkligen, egentligen, naturligtvis:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_essay_id&result_tab=2&search=cqp

Inte:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_essay_id&result_tab=2&search=cqp

Sökning 18 – Satsadverbial och enskilda informanter på olika kunskapsnivåer

Kanske, ju, verkligen, egentligen, naturligtvis:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_student_id,text_grade&result_tab=2&search=cqp

Inte:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22NA%22%20%26%20lex%20contains%20%22inte%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&stats_reduce=lex,text_student_id,text_grade&result_tab=2&search=cqp

Antalen informanter jämfördes med de totala antalen informanter på de olika nivåer som hittades via:

https://spraakbanken.gu.se/korp/#?cqp=%5B%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&stats_reduce=text_student_id,text_grade&search=cqp

Sökning 19 – Modala satsadverbial i skönlitterära texter

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=aspacsv,romi,romii,rom99,storsuc,romg&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

Sökning 20 – Modala satsadverbial i Twitter-texter

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=twitter,twitter-pldebatt-130612,twitter-pldebatt-131006,twitter-pldebatt-140504,twitter-2015,twitter-2016,twitter-2017&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

Sökning 21 – Modala satsadverbial i tidningstexter

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=gp1994,gp2001,gp2002,gp2003,gp2004,gp2005,gp2006,gp2007,gp2008,gp2009,gp2010,gp2011,gp2012,gp2013,gp2014,press65,press76,press95,press96,press97,press98,webbnyheter2001,webbnyheter2002,webbnyheter2003,webbnyheter2004,webbnyheter2005,webbnyheter2006,webbnyheter2007,webbnyheter2008,webbnyheter2009,webbnyheter2010,webbnyheter2011,webbnyheter2012,webbnyheter2013,attasidor,dn1987,ordat&search_tab=1&result_tab=4&stats_reduce=lex&search=cqp

Sökning 22 – Modala satsadverbial i akademiska texter

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%5D&corpus=sweachum,sweacsam&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

Sökning 23 – Modala satsadverbial och dependensrelationer (med orden kanske, ju, verkligen, egentligen och naturligtvis)

Inlärmaterialet:

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&stats_reduce=deprel&search=cqp

Talbanken:

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=talbanken&search_tab=1&result_tab=2&stats_reduce=deprel&search=cqp

Sökning 24 – Egentligen i Bloggmix-materialet

Vilken som helst dependens:

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22egentligen%5C.%5C.ab%5C.1%22%5D&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmix2017,bloggmixodat&search_tab=1&search=cqp

Som satsadverbial:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22egentligen%5C.%5C.ab%5C.1%22%5D&corpus=bloggmix1998,bloggmix1999,bloggmix2000,bloggmix2001,bloggmix2002,bloggmix2003,bloggmix2004,bloggmix2005,bloggmix2006,bloggmix2007,bloggmix2008,bloggmix2009,bloggmix2010,bloggmix2011,bloggmix2012,bloggmix2013,bloggmix2014,bloggmix2015,bloggmix2016,bloggmix2017,bloggmixodat&search_tab=1&search=cqp

Sökning 25 – Modala satsadverbial på nivå A1, kontrollsökning med kanske, ju, verkligen, egentligen och naturligtvis

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22egentligen%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&result_tab=2&global_filter=eyJ0ZXh0X2dyYWRIIjpbIkExIl19&search=word%7Ckanske&page=0

Sökning 26 – Modala satsadverbial på nivå C1, kontrollsökning med egentligen, ju, naturligtvis och verkligen

https://spraakbanken.gu.se/korp/#?cqp=%5Blex%20contains%20%22egentligen%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&page=0&global_filter=eyJ0ZXh0X2dyYWRIIjpbIkMyIl19&search=word%7Cegentligen

Sökning 27 – Modala satsadverbial i inlärttexter på nivå C1 (med kanske, ju, egentligen, enligt och verkligen)

Med dependensrelation satsadverbial:

https://spraakbanken.gu.se/korp/#?cqp=%5Bdeprel%20%3D%20%22MA%22%20%26%20lex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%20%26%20_.text_grade%20%3D%20%22C1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&search=cqp&stats_reduce=lex

Vilken som helst dependensrelation:

https://spraakbanken.gu.se/korp/#?cqp=%5B_.text_grade%20%3D%20%22C1%22%20%26%20lex%20contains%20%22kanske%5C.%5C.ab%5C.1%22%5D&corpus=spin-source,sw1203v2,tisusv2&search_tab=1&result_tab=2&stats_reduce=lex&search=cqp

BILAGA 2. Samtliga ord som taggats som satsadverbial på olika kunskapsnivåer i inläromaterialet.

A1	
<i>Ord</i>	<i>Träffar</i>
i	1
hem	1
\$\$ster	1

A2	
<i>Ord</i>	<i>Träffar</i>
kanske	11
även	1
ändå	1
visst ⁵⁴	1
ute	1
Tyvärr	1
pratade	1
mycke	1
med	1
komer	1
jo	1
Igår	1
hjalpte	1
förstås	1
bättre	1

B1	
<i>Ord</i>	<i>Träffar</i>
kanske	17
ju	9
Tyvärr	4
Naturligtvis	4
verkligen	3
tror	3
väldigt	2
tycker	2
Visst	1
säger	1
skulle	1
naturligtvis	1
gärna	1
Enligt	1
engenligen	1
egentligen	1
Däremot	1
Avslutningvis	1
Avslutadvis	1
anpassa	1
är	1
Äntligen	1
man	1
har	1
förstås	1
förrular	1

⁵⁴ Inte ett satsadverbial, utan en felstavad form av 'visste'

B2			
<i>Ord</i>	<i>Träffar</i>	<i>Ord</i>	<i>Träffar</i>
kanske	43	snarare	1
verkligen	17	Slutligtvis	1
ju	11	Sammanfattningsvis	1
förstås	6	ont	1
egentligen	6	Också	1
Kanske	6	mer	1
Avslutningsvis	5	igen	1
naturligtvis	4	göra	1
säkert	4	ganska	1
Naturligtvis	4	förfarande	1
faktiskt	4	förstås	1
väl	3	finns	1
Visst	3	Emellertid	1
givetvis	3	Dock	1
tror	3	bero	1
nog	3	ändå	1
gärna	3	väsentligt	1
enligt	3	tro	1
tycker	2	som	1
Slutligen	2	precis	1
sannolikt	2	ORKA	1
I	2	långt	1
eventuellt	2	långsammare	1
väldigt	1	hitta	1
var	1	gå	1
utan	1	Följaktligen	1
Troligen	1	anstränga	1
tillräckligt	1		

C1			
<i>Ord</i>	<i>Träffar</i>	<i>Ord</i>	<i>Träffar</i>
kanske	49	ända	1
ju	27	Än	1
egentligen	21	återigen	1
verkligen	18	Ytterligare	1
enligt	15	vara	1
säkert	9	tycker	1
nog	8	säger	1
Naturligtvis	6	säga	1
Tyvärr	6	således	1
förstås	6	styrk	1
naturligtvis	6	stressat	1
väl	5	självklart	1
Visst	5	Redan	1
Enligt	5	Plötsligt	1
Sammanfattningsvis	4	nödvändigtvis	1
helt	4	nästän	1
Kanske	4	något	1
faktiskt	4	Nuförtiden	1
helst	3	möjligen	1
antagligen	3	mycket	1
som	2	lära	1
givetvis	2	lite	1
tyvärr	2	kräver	1
tror	2	jobba	1
mer	2	Jo	1
gärna	2	Inte	1
Givetvis	2	I	1
Dessvärre	2	har	1
är	1	förvisso	1
väldigt	1	förmodligen	1
troligen	1	Egentligen	1
kommer	1	dåligt	1
Idag	1	bo	1
Avslutningsvis	1	Arbetar	1
överhuvudtaget	1	absolut	1

C2			
<i>Ord</i>	<i>Träffar</i>	<i>Ord</i>	<i>Träffar</i>
kanske	9	nog	1
verkligen	2	naturligtvis	1
Tyvärr	2	Kanske	1
överhuvudtaget	1	ju	1
tror	1	förstås	1
säkert	1		