

GÖTEBORGS UNIVERSITET  
Institutionen för svenska, flerspråkighet och språkteknologi

Expletivt *det* som formellt subjekt i texter skrivna av  
andraspråksinlärare

Håkan Gunnarsson

Kandidatuppsats, SSA 136 15 hp  
Svenska som andraspråk  
HT 2022  
Handledare: Julia Prentice

## Sammandrag

Syftet med studien är att utifrån inläratexter som ingår i L2-korpusar som förvaltas av Språkdata vid Göteborgs universitet studera bruket av expletivt *det* som formellt subjekt.

Det är en företrädevis kvantitativ studie som drar nytta av de taggningar och i vissa fall korrigerande annoteringar som gjorts, vilket inte bara tillåtit frekvensundersökningar med avseende på språklig färdighetsnivå och modersmål, utan också felanalyser med avseende på dessa parametrar.

Jämförelser med förstaspråksanvändare (L1) har gjorts, i med hjälp av tidigare publicerade studier om expletivt *det* och närliggande fenomen, men även egna korpussökningar i L1-korpusar.

Resultaten visar i stort att frekvenser av konstruktionen i inläratexten (L2) ligger nära användningen av förstaspråksanvändare (L1) eller har avvikelser av förklarliga skäl, samt en förväntad utveckling i användningen vad gäller färdighetsnivå.

Nyckelord: *andraspråksinlärning, formellt subjekt, expletivt "det", korpus, L2-svenska, helfraser.*

# Innehållsförteckning

1. Inledning – 1
2. Syfte – 2
3. Teori och tidigare forskning – 3
  - 3.1. Teoretisk ram – 3
    - 3.1.1. Input och output – 3
    - 3.1.2. Implicit och explicit – 4
    - 3.1.3. Talat och skrivet – 6
    - 3.1.4. Om felanalys: i korthet – 8
    - 3.1.5. Tidigare forskning – 8
4. Bakgrund: Expletivt *det* – 10
  - 4.1. Presentering – 11
  - 4.2. Postponerat led – 12
  - 4.3. Expletivt *det* med *med*-fras – 13
  - 4.4. Utbrytning – 13
  - 4.5. Expletivt *det* med passiv konstruktion – 14
5. Metod – 15
6. Material – 17
  - 6.1. Genrer – 19
7. Resultat och diskussion – 20
  - 7.1. Förekomster – 20
    - 7.1.1. Utbrytningskonstruktionen – 20
    - 7.1.2. Expletivt *det* med predikativt attribut och *med*-fras – 23
    - 7.1.3. Expletivt *det* med presentering och postponerat led – 24
    - 7.1.4. Expletivt *det* och passiv konstruktion – 25
  - 7.2. Felanalys: Expletivt *det* som formellt subjekt och fokusfel –  
utbrytningskonstruktionen – 25
    - 7.2.1. Expletivt *det* som formellt subjekt – 25
    - 7.2.2. Expletivt *det* i presentering och postponerat led och inlärares  
modersmål – 28

- 7.2.3. Utbrytningskonstruktionen – fokusfel – 29
- 7.3. Fördjupning: *det* och *det finns* som högfrekventa helfraser 30
- 8. Avslutning – 33
  - 8.1. Slutsats – 33
  - 8.2. Vidare forskning – 34
- 9. Litteratur – 36

# 1. Inledning

Språk handlar om överföring av information. Det är möjligen den bredaste inramningen av den här uppsatsens undersökning. Givet den ramen kan det uppfattas som motsägelsefullt att intressera sig för något så innehållslöst som ett *det* som saknar referens. Men med tanke på hur vanliga subjekt av den här typen är i svenska språket och att språket rimligtvis i varje situation försöker vara så effektivt som möjligt, finns det uppenbarligen en underliggande komplexitet som är värd att undersöka, inte minst ur ett andraspråksperspektiv. Om det finns någon väsentlig företeelse förknippat med ett visst språk, har detta relevans även för inläraren av ett främmande språk.

Expletivt *det* (hädanefter även: ED) är representerat i språket både som närmast mekaniskt och vanemässigt använda helfraser som en tematisk-informativ resurs som kanske få utanför modernmålstalare riktigt behärskar.

*Det* är det vanligaste ordet i svenskan, en ställning ordet inte bara fått genom funktionen att vara ett pronomen för neutrala substantiv utan som ett slags bestämd artikel och dess roll just som icke-referentiellt pronomen och platshållare i funktion för subjektstvånget i svenskan. Enligt korpusstudier (refererade till i Forsskåhl 2009: 288; Engdahl 2010: 97) är vidare *det är* det vanligaste ordparet i svensk skriven text och samtal.

Givet detta faktum är det ur ett andraspråksperspektiv ofrånkomligt att inlärare tidigt kommer i kontakt med denna konstruktion, och tanken med den här uppsatsen är att bidra till kunskaper om dess användning.

Det expletiva pronomenet *det* fungerar som subjekt i olika satstyper. I de flesta fallen verkar konstruktionen vara ett sätt att ge *det* led som annars skulle ha fungerat som subjekt vid satsens predikat en plats i satsens tematiska del (SAG 4: 53).

I *Svensk ordbok* knyts *det* som pronomen (SO: ”det”) till (1) subjekt i opersonliga konstruktioner, (2) funktion som formellt subjekt, (3) konstruktion med predikativ och emfatisk utbrytning.

Svenska Akademiens grammatik (SAG 4: 53–63) viger sammanlagt tio sidor åt expletivt *subjekt*, och utför samma grundläggande indelning.

Då ED innebär ett subjekt *utan* information ger detta en speciell funktion åt konstruktionen. Begreppsparet *tema* och *rema* används bland annat inom textanalysen inom vilken man vanligen lånar termer av den systematisk-funktionella grammatiken. Melin och Lange (2000: 76) förklarar distinktionen med att de respektive termerna betecknar plats i meningen och den information som förmedlas värderas med avseende på känd respektive ny information.

Eftersom den här uppsatsen rör sig på meningsnivå, kommer inte några vidare analyser kring tema och rema att göras, utan anledningen att nämna det här, är för att sätta in begreppen så som de förekommer i beskrivning av funktionen hos ED i ett större betydelse- och användningssammanhang.

## 2. Syfte

Efter den här introduktionen till satsgrammatiska konstruktioner med expletivt *det* som formellt subjekt, behöver undersökningens syfte preciseras.

Syftet är att undersöka användningen av expletivt *det* som formellt subjekt i inlärttexter. Följande forskningsfrågor ska hjälpa till att kasta ljus över denna användning.

- Vilka konstruktioner med expletivt *det* som formellt subjekt förekommer och i vilken utsträckning med hänsyn tagen till kunskapsnivå?
- Vad kan sägas, om möjligt, angående förekomsten i inlärttexterna jämfört texter skrivna av förstaspråkstalare?
- Vilka avvikelser med just detta subjekt kan man finna jämfört med andra subjekt som tyder på svårigheter att använda dem?

- Finns det skillnader mellan hur olika språkbakgrund påverkar förmågan att använda expletivt *det*?

## 3. Teori och tidigare forskning

### 3.1. Teoretisk ram

För att få en överblick över problematiken om ED följer här en genomgång av ett antal perspektiv och metoder som används inom forskningen, både inom svenska, svenska som andraspråk och lingvistik.

#### 3.1.1. *Input och output*

För att ringa in det teoretiska underlaget för uppsatsen är det lämpligt att först relatera till den klassiska dikotomin, mellan modeller som utgår från betydelsen av språkligt intag, *input*, och den språkliga produktionen, *output*. Det handlar också om frågan om det spontana i förhållande till det planerade, och *uppmärksamheten* i muntlig språkproduktion. Säkert alla SFI-lärare har varit med om elever som offrat flyt i talet till förmån för korrekthet och tvärtom. Även om den distinkta dikotomin de senaste decennierna blivit mindre aktuell kan den ändå tjäna till att förtydliga två perspektiv. I enlighet med Krashens monitormodell, låter den noggranna eleven sin explicita kompetens om språket, som eleven tillägnat sig spontant genom tillgång till ett inflöde, kontrollera utflödet med avseende på korrekthet. Men den egentliga kompetensen är den som sker utan ”monitor”; den når i slutänden det bästa kommunikativa syftet (Abrahamsson 2009: 119). Det är till och med så att Krashen avvisar att medvetandet har någon roll i inlärning av språket; det är en undermedveten process (Ellis 2008: 265). Med Krashens utgångspunkt kan vi sluta oss till att vadhelst för kvalitéer det är eleven visar upp i texter, är det en kompetens som byggts av en input av rikligt innehåll, snarare än form; det är bara det att monitorn får en mer självklar plats i textproduktion.

Krashens fokus på input kontrasteras i vetenskapsöversiktarna mot Merrill Swains betoning av den muntliga *produktionens*, ”outputens”, betydelse för att stärka interimsspråket. Inläraren blir i språklig interaktion med andra personer tvingad att formulera sig begripligt genom att med grammatiska medel analysera utsagorna, och vad det var som eventuellt gjorde dem oförståeliga för lyssnaren (Abrahamsson 2009: 178f.). I den här kognitivt inriktade teoretiska modellen hos Swain bidrar både en ökad automatisering av regelföljande *och* uttryck som oanalyserade sammanhängande block till behärskning av målspråket (Saville-Troike & Barto 2017: 80).

I termer av ”input” och ”output” på kognitiv grund har mer elaborerade teorier utvecklats de senaste tjugo-trettio åren. En av de mer framträdande är *konnektivismen*. Den kan sägas höra till de teorier som bygger på användning, ”usage-based”, det vill säga föreställningen att språktillägnande bygger på den praktiska användningen av målspråket. I fallet med konnektivismen innebär det specifikt att inlärarna använder sig av generella kognitiva mekanismer – inte exklusiva för språkinlärning – för att bilda reglerna för sitt interimsspråk. Det handlar vidare om associativt lärande av form, innehåll och funktion, och om en rationell kognitiv bearbetning av dessa vilken är beroende av frekvens och språkligt sammanhang. Mycket av detta är ändå implicit, omedvetet. Hela denna inlärningsprocess, enligt konnektivismen, är en gradvis process där målspråket utvecklas både som ett komplext och anpassningsbart system som hela tiden korrigeras i detaljerna genom att de beskrivna enkla kognitiva mekanismerna interagerar med andra språkanvändare i skilda sociala sammanhang (Ellis & Wulff 2020: 63f.).

### ***3.1.2. Implicit och explicit***

En viktig distinktion inom den kognitiva andraspråksforskningen har varit den mellan implicit och explicit kunskap, där implicit kunskap inte är beroende av att medvetet ta minnet i bruk.

Eftersom texterna som utgör källorna till den här uppsatsen skrevs av informanter som var i ett sammanhang av utbildning i svenska, eller möjligen till en urvalsprocess till sådan utbildning, kan man på goda grunder anta att de stått under inflytande av explicit undervisning.



Vid närmare betraktande är det inte en helt lätt distinktion att definiera och upprätthålla, även om det för Krashen är en principiell viktig dikotomi. Men var går till exempel gränsen mellan explicit kunskap som blir så automatiserad att den *funktionellt* inte går att skilja från implicit (Ellis 2008: 418f.)?

Ett sätt att närma sig frågan om implicit respektive explicit kunskap, som utförts redan på 70- och 80-talen, är att mäta det förstnämnda som de kunskaper inläraren faktiskt uppvisar i muntlig och skriftlig produktion och det sistnämnda genom de förklaringar som inlärarna kan leverera på språkliga fenomen (Ellis 2008: 424ff). Redan här kan man resa frågan hur man skulle kunna säkerställa att den implicita kunskaper verkligen är rent implicit, speciellt om man tänker på de inlärare som verkligen tänker efter innan de talar och väljer att offra flyt mot korrekthet.

En närliggande distinktion till implicit och explicit kunskap är den som beskriver språklig utveckling i termer av pragmatisk eller grammatisk utveckling. Frågan är vilken av de båda senare som kommer först och hur de förhåller sig till varandra. Sambanden har studerats i olika longitudinella studier som pekar i riktning mot att en pragmatisk språklig kompetens föregår grammatisk, men här spelar också kontexten för inläring in (Ellis 2008: 195ff.). För den grupp som Ellis (2008: *ibid.*) kallar för "naturalistic learners" föregår den pragmatiska kompetensen den lingvistiska i de tidiga faserna, men det krävs grammatisk utveckling för att klara av de mer sofistikerade nyanserna i även en pragmatisk användning. Och på den andra extremen, där det grammatiska kan föregå det pragmatiska, har vi inläring av *främmande* språk, det vill säga när inläraren inte lever i målspråkets sociala kontext (Ellis 2008: *ibid.*).

När det gäller explicit *grammatikundervisning* är det rimligt att anta att den är inriktad på språkriktighet, oaktat hur återkoppling sker kvantitativt och kvalitativt. Om det dessutom gäller välutbildade informanter är det lika rimligt att anta att de sört för sin egen inläring om de inte stått under organiserad undervisning. Eftersom bara vissa av ED-konstruktionerna undervisas explicit över huvud taget – likt så många andra språkliga fenomen i svenskan – finns det skäl att anta att de lärt sig dessa genom passivt tillägnande, till exempel genom textläsning och inte minst talad svenska, vardaglig interaktion såväl som vanligt berättande.

Enligt Doughty (2003: 259ff.) finns det metodologiska svårigheter inom andraspråksforskningen med att studera effekterna av lärande ("learning") i meningen

undervisning ("instruction") och dess effekt på tillägnande ("acquisition"). Distinktionen 'lärande' och 'tillägnande' går förstås tillbaka på Krashens teorier som refererats till ovan. Det finns en motsättning mellan att få en så kontrollerad miljö som möjligt men samtidigt ta hänsyn till att det är i verkliga klassrum, inte laboratorier, som undervisningen sker. Williams (2012: 549) menar sig kunna bekräfta den allmänna uppfattningen att explicit klassrumsundervisning faktiskt gör skillnad och tillägger att forskningen de senaste par decennierna lyfter betydelsen av såväl samtal och interaktion som komplexa och "metalingvistiska" kunskaper – det senare något kontroversiellt i interaktions- och kommunikationsbetonad undervisning.

Det finns en avgörande skillnad mellan att till exempel studera inlärarproduktion med avseende på satsadverbialets placering före det finita verbet i bisats, jämfört med inlärarproduktionen av ED, i det att ED ofta är optionellt, och utebliven ED-sats inte blir ogrammatisk om än möjligen icke-idiomatisk: Den som vill negera en proposition måste placera satsadverbialet *inte* på rätt plats i satsen. Som jämförelse blir tillägandet av ED på en nivå som korresponderar med L1-kompetens något som tagit andra vägar än explicit undervisning (om syntax). Samtidigt blir inte konstruktionen mindre intressant utifrån ett L2-perspektiv.

När det gäller expletivt *det* och dess användning, kan en hypotes vara att om den förekommer i inlärarpråk har den tillägnats på ett implicit sätt. Enligt egen erfarenhet av ett dussintal SFI-läroböcker, däribland de frekvent använda Scherrer & Lindemalm (2014, 2015), Fasth & Kannermark (1996, 1997, 1998) och Ballardini et al (1986), är det inte ens så att den undervisas om explicit på det sätt man explicit undervisar om till exempel ordföljd eller tempusföljd.

### 3.1.3. *Talat och skrivet*

Om vissa typer av konstruktioner med ED är mer frekventa och mångsidigt använda i det talade språket, påkallar den här studien som behandlar skriftlig produktion några fundamentala skillnader mellan talat och skrivet. Det skriftliga är mer strukturellt stabilt och normgivande för grammatiken, medan det talade uppvisar större variation och flexibilitet än det talade; i talet finns ett sampel och turtagning som tillsammans

med ämnesutveckling och initiativtagande utgör talets *dialogiska* organisation (Lindström 2008: 27f.).

Förhållandet mellan skriven produktion och muntlig produktion är komplex. Även om man gärna förknippar inläring av L2 med någon form av lärsituation i en skolkontext med läroböcker, läsning och skrivande, har de allra flesta teorier och modeller om L2-inläring och L2-kompetens formulerats för *muntlig* produktion (Polio 2012a: 320). Polio (2012a: *ibid*) menar till och med att gängse teorier kring andraspråksinläring inte ens är direkt applicerbara på skrivande, och de som försökt klassificera dem gör det alla lite olika (beroende på att teorierna överlappar och inte alltid är ömsesidigt uteslutande).

Kopplingen mellan explicit och implicit kunskap och skrivande är, som Polio (2012a: *ibid*) framhåller, att vid skrivande har inläraren tid att bedöma och korrigera vad som skrivits och på så sätt sätta använda sig av inlärd regler. Så varje modell för L2-skrivande måste ta ställning till vilken roll implicit respektive explicit kunskap spelar. För det är samtidigt så att inlärare i sin produktion inte alltid lyckas realisera sina kunskaper om regler i sina texter. Till detta kommer frågan hur arbetsminnet fungerar i tal respektive skrift (Polio 2012a: 322). Men det är till exempel inte självklart att textskrivande blir mer grammatiskt korrekt och idiomatiskt bara för att eleven har mer tid på sig (Makalela 2004).

Sammanställningen av sex olika teorier (eller "approaches") av Polio (2012b: 378) är nedslående för den skrivna produktionens relevans som mått på tillägnande av ett andraspråk. Det är egentligen bara bland sociokulturella ingångar som skriven text uppmärksammas på allvar och genererat studier att tala om. Den ovan, under konnektivismen (3.1.1), refererade ingången till andraspråksinläring och kompetens som handlade om "usage-based", hör till dem som har inget eller svagt intresse för skriven produktion, och fram till tidpunkten för artikelns tillkomst har inga studier gjorts.

Forskning om skrivande och vuxna inlärare av L2-engelska uppvisar en övervikt av högpresterande inlärare (Gonzalves 2019: 7). Till studier med den sociolingvistiska anknytningen av svenska inlärare och skrivande kan nämnas Winlunds (2021) studie av ungdomar och skrivande i klassrumssituationer.

### 3.1.4. *Om felanalys: i korthet*

Felanalysen har i andraspråksforskningen en lång historia av olika teoretiska motiveringar och praktiska tillämpningar, där teorier med ambitioner på förutsägelser och finlandet av enkla orsakssamband har förlorat terräng utifrån de empiriska studier på den faktiska L2-användningen som visat på komplexiteten bakom varför målspråksinlärare gör de fel de gör. Den här studien behandlar ett syntaktiskt fenomen, och när det gäller den långa tradition av att undersöka i vad mån och hur modersmålet spelar roll för de fel som görs, framhåller Abrahamsson (2009: 243) att modersmålet har klart mindre inflytande på morfologi och syntax än till exempel på det fonologiska och lexikala. När det gäller fenomenet ”markerad struktur” som bidrog till att utveckla och nyansera felanalysen, sammanfattar Abrahamsson (2009: 245) en komplicerad orsaksbild med att ökat markeringsskillnader mellan L1 och L2 ”tenderar [inlärare] att ersätta högt markerade drag i målspråket med lågt markerade drag”.

### 3.1.5. *Tidigare forskning*

Med Thyberg (2020) som undantag i nutid har studier på ED, och relaterade konstruktioner, till övervägande del analyserat muntligt material. Om man ser till den möjligen främsta anledningen till konstruktionen, nämligen att den viktiga satsinformationen, vanligtvis den som är *ny*, ska komma sent i satsen, är det inte konstigt att detta språkliga grepp framför allt används i tal där åhöraren ska uppfatta informationen i realtid, inte som vid läsning av text där möjlighet till omtag och val av lästempo finns (jfr. Ekroth 2011: 106). Detta hindrar dock inte att en text har potential att bli mer läsvänlig om samma konstruktioner används i skrift.

Expletivt *det*:s höga frekvens i förekomst leder oss in på ett specifikt spår i undersökningen. I ett andraspråksperspektiv har det bedrivits studier med hypotesen att vid inläring blir vissa fraser vilka formellt består av subjekt och verb – däribland *det är* – en odelbar helfras som alltid lagras i minnet och används som denna odelbara helhet. Axelsson (1989) har i ett muntligt material bestående av inlärtal i nybörjarstadiet sett hur *de' e* -fraser följt av nominal- eller adjektivfraser är mycket

frekventa samt hur dessa under inlärningsprogressionen kompletteras med fler varianter på satsavslut under det att helfraserna förblir få; det handlar fr.a. om *de' e*-fraser som kompletteras med *de' måste* och vidareutvecklas med olika personliga pronomen. De här helfraserna, menar Axelsson (1989: 199), ger inläraren möjlighet att hålla en konversation utan att behöva bearbeta språkregler eller vinna viss tid att skapa resten av yttrandet. Samtidigt ansluter inläraren till en extremt frekvent och därmed idiomatisk användning av existensutsagor i svenskan. Hammarberg (2008) följer upp detta spår med helfras-användning, och replikerar Axelssons arbete på ett muntligt material och anknyter även till senare internationell forskning som uppmärksammat *frekvensens* roll i språkinläringen. Där blir hela strukturer, till exempel *de' e*, befästa som ”kognitiva rutiner” som används och utvecklas vidare genom upprepad användning vid specifika tillfällen. Denna forskningsinriktning betraktar alltså språkinläring utifrån ett *användningsperspektiv* (engelska: ”Usage-based model”). På basis av erfarenheten hos inläraren blir språkliga enheter befästa genom mentala representationer. Hammarberg talar här om en användning som faller under beteckningen ”formelsekvens” (engelska: ”formulaic sequence”). Termer som figurerar i det här sammanhanget är ”chunks” och ”prefabs” vilket ger en bild av helhets-karaktären hos dessa fraser, eller med Hammerbergs ord, *sekvenser* (jfr. Ellis 2008: 75ff.). Utifrån detta borde man kunna tänka sig att även andra expletiva *det* rutiniserar på samma sätt som en helhet, sammanhängande block som lagras som sådana i minnet. Fördelen med dessa ”chunks” är att de är ekonomiska ur ett bearbetnings- eller analysperspektiv, samtidigt som de står i en relation till en mera krävande kreativ process av produktion, eller avkodning, av mera komplexa yttranden (Hammarberg 2008: 99f; jfr. Bolander 1989).

En del av forskningsinriktningen utifrån användningsperspektiv har en psyko- eller neurolingvistisk karaktär, och begränsar sig inte till de fraser eller muntliga satser som har den helhetskaraktär som Axelsson (1989) formulerade (Winlund 2021: 2).

Thyberg (2020) samlar sitt material från skriftliga korpusar bestående av L1-svenska, och studerar de nämnda konstruktionerna utifrån systemisk-funktionell grammatisk teori och terminologi, för att kartlägga hur icke-referentiellt *det* konstruerar betydelse bortom de funktioner som har med syntax (till exempel plastshållare) och informationsstruktur att göra. Följaktligen har inte fokus varit att

till exempel undersöka högerpositioneringen av rematiska led (Thyberg 2020: 22, 260).

Ekerot (1979) tar upp talat och skrivet när gäller olika transformationer av satser, som ED-konstruktioner är exempel på, och konstaterar att vissa är talspråkligt präglade och andra skriftspråkligt. Till de förra räknar han utbrytningskonstruktionen (se nedan) och till de senare adverbialförflyttning. Allt handlar om att givet produktionssättet, muntligt eller skriftligt, ”signalera informationsstruktur” (Ekerot 1979 104f.).

Forsskåhl (2009) behandlar *det är* som ”resurs” i muntlig kommunikation, och håller öppet om hennes observationer även gäller skriftlig produktion (Forsskåhl 2009: 263f.). Den studien är dock baserad på analys av samtal med noggrann uppmärksamhet på turtagning mellan deltagare, och detta kan sägas ligga en bit från eventuella användningar av *det är* i de olika essäuppgifter vi kommer att använda oss av (se nedan, avsnitt 6), oaktat att det inte heller handlar om samtal mellan inlärare av svenska. Samtidigt är studier av ED av relevans för andraspråksinlärare, eftersom man borde kunna anta att språket runtomkring dem, inklusive skrivna texter, påverkar deras användning av konstruktionen i varierande grad.

## 4. Bakgrund: Expletivt *det*

Av alla de konstruktioner där ED förekommer, kommer vi beröra ett urval av de mest frekventa eller, de SAG menar att ED ”används särskilt” inom (SAG 4:53). SAG talar i sin genomgång om ”expletivt *subjekt*”, men de konstruktioner som den här uppsatsen behandlar har alla *det* som detta subjekt.

ED ingår i ett sammanhang av språklig variation. När vi möter ED i text (och tal) handlar det om ett val av konstruktion för ett syfte. Som Josefsson (2009: 209) konstaterar hänger valet ihop med vilka i konstruktionen som fokuseras och vilka delar som kommer i bakgrunden, ”hur informationen ’packas’” (Josefsson 2009: *ibid*).

Ett begrepp som förekommer i sammanhanget av ED är *existentialsatser*, till exempel i Sundmans (1980) inflytelserika studie om subjektproblematiken i svenskan. Där räknas till exempel inte opersonliga subjekt till existentialsatser.

Existentialsatser ska innehålla såväl formellt som egentligt subjekt (Sundman refererad i Hammarberg 2008: 89)

I Svenska Akademiens grammatiks (SAG 4: 53; SAG 1: 168) redogörelse för ED särskiljs det fall då ED används vid opersonligt predikat, till exempel det vanliga sättet att beskriva väderförhållanden som inlärare av svenska L2 i undervisning möter tidigt:

(1) Det snöar.

Den konstruktionen kommer inte särskilt intressera oss i den här uppsatsen. Detta eftersom vi avgränsar oss från ED i konstruktioner där man från början kan förvänta sig att de lärts in som hela uttryck, och att det inte finns något annat sätt att uttrycka det, likt andra fraser som handlar om upplevda omständigheter.

#### 4.1. Presentering

Av intresse är i stället de satser som innehåller egentligt subjekt ED som *det* som formellt subjekt, alltså där det enda syftet synes vara att satsens subjekt eller predikationsbas flyttas som beskrivet ovan, vilket sker i sammanhanget av att introducera ny information. Typiskt förekommer ett rumsadverbial i satsen (Josefsson 2009: 143). Exempel på detta är när bara själva existensen utsägs, respektive när något utsägs om detsamma:

(2) Det finns mjölk i kylskåpet.

Som tidigare nämnts är det ofta men *inte alltid* det egentliga subjektet förekommer i satsens rematiska del, när ED uppträder med egentligt subjekt i en sats. Detta egentliga subjekt kan inte vara bestämt. I närvaro av andra satsled är det inte nödvändigt att remat kommer sist, bara att det inte kommer först (Ekerot 2011: 106).

Omöjligheten att ha egentligt subjekt i bestämd form kan vara ett sätt att spåra felaktigt inlärda presentationskonstruktioner.

Det kan uppfattas som onödigt ur ett ordekonomiskt perspektiv med en satskonstruktion innehållande ED och egentligt subjekt, eftersom konstruktionen inte

tillför någon ytterligare information i kvantitativ mening, även om alltså placeringen av det egentliga subjektet i rema-delen fyller en annan funktion. I det ovanliga fallet där det egentliga subjektet står på fundamentplats framstår ED som ännu mer överflödigt. Jämför följande två meningar:

- (3) Avloppsvatten väljer det återigen upp ur golvbrunnen.
- (4) Avloppsvatten väljer återigen upp ur golvbrunnen.

Frågan är vad det är som framhävs i vart och ett av dessa två satser. Den första kan möjligen framhäva ”uppvällandet” relativt det likaledes framhävda ”avloppsvattnet” eller satsadverbialiet ”återigen”; medan den andra satsen, genom placeringen av substantivet som subjekt, utgör mera bara ett neutralt konstaterande? Eller är det alltför svårt att fastställa en betydelskillnad i någon som helst bemärkelse?

I föregående fall med ett egentligt subjekt som fundament, kan detta egentliga subjekt inte utgöras av en bisats eller postponerad infinitivfras (SAG 4: 55f.).

#### 4.2. Postponerat led

En annan konstruktion där ED kommer i funktion är i kombination med ett *postponerat* led. Till skillnad mot ED konstruktioner med presentering, *med*-fras (se nedan) och utbrytning där konstruktionen med ED är högst valfri, är den här i många fall konstruktionen starkt att föredra eller helt nödvändig. Detta postponerade led kan vara en infinitivfras, en bisats med påstående, indirekt fråga eller utrop (narrativ, interrogativ, expressiv). När dessa står som fundament kan meningarna konstrueras utan ED (SAG 4: 55f.).

- (5) Att undvika höjda elpriser i vinter kommer bli svårt.
- (6) Det kommer bli svårt att undvika höjda elpriser i vinter.
- (7) Det är otroligt vad otrevlig han låter!
- (8) Vad otrevlig han låter är otroligt!



Bland flera funktioner för konstruktionen med postponerat led, vid sidan om det gängse att låta ny information komma senare i meningen, nämner Thyberg (2020: 34) att den ger möjlighet att ge ett omdöme om informationen i ledet.

#### 4.3. Expletivt *det* med *med-fras*

En annan konstruktionstyp med ED som SAG lyfter fram är där det egentliga subjektet står som rektant till en *med-fras*. Denna konstruktion används ofta för att uttrycka värdering och samtidigt – likt övriga konstruktioner – förskjuta satsens tema till den rematiska delen där den får en kraftigt betonad position.

- (9) Som säkerhetsåtgärd kan det vara bra med extra kontroller.
  - (10) Just nu blir det utmärkt med en paus.
- Utan *med-fras*:
- (11) Extra kontroller kan vara bra som säkerhetsåtgärd.
  - (12) En paus blir utmärkt just nu.

#### 4.4. Utbrytning

En för inläraresammanhang känd konstruktion inom andraspråksundervisningen är den så kallade *utbrytningskonstruktionen* (eller: *emfatisk omskrivning* eller *satsklyvning*). Till skillnad från de övriga konstruktionstyperna av ED tas den vanligen upp i andraspråkläroböckerna, till exempel de välkända Scherrer & Lindemalm (2014, 2015), Fasth & Kannermark (1996, 1997, 1998) och Ballardini et al (1986) nämnda ovan (3.1.2). Utbrytningskonstruktionen syftar till att *fokusera* (det egentliga) subjektet, objektet eller ett adverbial (Ekerot 2011: 71). Subjektet lyfts ut ur satsen och blir egentligt subjekt i en existensutsaga med en relativsats fungerande som satsattribut.

- (13) Det är barnen som ska tömma diskmaskinen den här veckan.
- (14) Det är diskmaskinen som barnen ska tömma den här veckan

(15) Det är den här veckan som barnen ska tömma diskmaskinen.

Betydelseskilnaden mellan de tre meningarna ovan är här tydlig. I tal skulle det ur prosodisk synvinkel bland annat innebära att ”barnen” kan få en tydlig betoning, och ur tematisk synvinkel hur subjektet hamnar på satsens rematiska del.

I det här sammanhanget är det lämpligt att introducera de *sats-* och *textgrammatiska* problem som uppstår med svenskans regler för ordföljd som bestäms av den förstnämnda grammatiken. Satsgrammatiken uttrycker de grammatiska relationerna inom den individuella satsen, medan textgrammatiken ser funktionellt på delarna för att klargöra vad som är given respektive ny information (Ekerot 2011: 66). Inom textgrammatiken kan man tala om en tematisk och en fokuserande funktion där den tematiska är knuten till informationsaspekten. I språk med utvecklad kausböjning, som till exempel skiljer subjekt och objekt från varandra, finns en större frihet att på ett ekonomiskt och pregnant sätt placera den rematiska delen senare i satsen, samt att fokusera ett led (som också ensam som känd information i satsen) genom att spetsställa det. I svenskan kan man till exempel inte fokusera subjektet i frågor (Ekerot 2011: 69–73).

För andraspråksinlärare kan skillnader i textgrammatiken, alltså hur tema och rema placeras i satsen, skapa bekymmer. Det handlar om att informationsflödet blir stört i en text som annars satsgrammatiskt innehåller få och små fel (Ekerot 2011: *ibid.*).

Relevansen av detta för denna uppsats är att ge en överblick över vilka möjligheter och begränsningar som ges i svenskan med avseende på svenskans satsgrammatiska regler.

#### **4.5. Expletivt *det* med passiv konstruktion**

Med *passiv konstruktion* kan man med ED uttrycka underförstådd mänsklig agens.

(16) Det arbetades hårt fram till första pausen.

Även passiv konstruktion kan ha egentligt subjekt, ED som formellt subjekt, och kan enkelt omvandlas till sats utan ED med bibehållande av betydelse, precis som aktiva satser med egentligt och formellt subjekt.

(17) *Gauloises* röktes mycket på sjuttioalet.

Med ED:

(18) Det röktes mycket *Gauloises* på sjuttioalet.

(19) På sjuttioalet röktes [det] mycket *Gauloises*.

Som nämndes i inledningen är val av (sats)konstruktion ofta ett val som handlar om att betona, lyfta fram (jfr. Josefsson 2009: 209). Med passiv konstruktion har man till exempel valet att sätta ut agent eller inte göra det; gör man det inte sker vad Sundman (1987: 347) kallar en *avsubjektivering* av rollen *agent*, passivkonstruktionen hindrar agenten från att realiseras som subjekt. I den passiva konstruktionen förflyttas – med Sundmans ord – rollen som subjekt från ”den kontrollerande participanten” till en annan, alltså objektet i satsen och den aktiva satsens subjekt blir ett agentsadverbial med *av* (Sundman 1987: 350; 353). I egenskap av endast agentsadverbial kan det strykas helt (Ekerot 2011: 198). I många språk används passiv endast med anledning av att agenten är okänd eller att man inte omnämna den (Sundman 1987: 358). Tematiskt sker genom passiva konstruktioner en omkastning av satsleden, vilket gör det möjligt att använda denna konstruktion för att göra tema av ett satsled, som *inte* är subjekt, till ett subjekt (Ekroth 2011: 101).

## 5. Metod

Studiet av expletivt *det* som formellt subjekt har utförts med kvantitativ metod genom att ur korpustexter få fram förekomster och därmed frekvenser.

För att finna de eftersökta konstruktionerna med ED i inlärtexterna har jag använt mig av Korp, sökverktyget för Språkdatas textdatabaser. Att söka på syntaktiska företeelser är en utmaning eftersom man vill konstruera sökningarna på ett sådant sätt att man inte missar konstruktioner på grund av begränsningar i själva

sökkonstruktionen. Samtidigt vill man inte få alltför många ovidkommande sökresultat. I mån av resurser kan man göra breda sökningar som sedan manuellt rensas från dessa oönskade sökträffar. Å andra sidan kan man då inte dra nytta de i samband med sökningarna mycket användbara sammanställningar på olika parametrar som ord, modersmål, kunskapsnivå, med flera.

En anledning till att vara relativt utförlig vid återgivandet av själva sökningarna i Korp är att skapa den transparens kring resultaten som behövs för att kunna replikera eller framför allt förbättra resultaten i en större studie.

Svenska språkets stränga regler för placeringen av subjekt, verb och eventuellt satsadverbial, gör att sökningar på för frågeställningen relevanta och regelrätta konstruktioner borde kunna göras effektiva. En av poängerna med ett formellt subjekt är just att den står på den bestämda platsen för subjekt, vilket i sin tur är styrt av verbet vars placering har prioritet. Språkanvändaren har stor valfrihet att placera olika led som fundament, men när väl det valet är gjort, finns liten möjlighet till variation av placering av leden. Däremot är det en komplikation vid komposition av söksträngar att leden kan vara utbyggda med flera ord, och att det därför för varje tänkt satsled måste finnas utrymme för flera ord. Om man i sökningarna tillåter flera ord per satsled innebär det – som nämnts ovan – samtidigt att det kommer med ovidkommande sökträffar.

Sökningar på *det* som subjekt kommer att fånga in referentiella, icke-expletiva *det*, och därför är det värdefullt att kunna använda korpusar i *SweLL* som är taggade med avseende på typ av satsled, vilket för uppsatsens vidkommande innebär formellt subjekt. Givet den här studiens begränsade omfattning kommer det som i materialet och av sökmaskinen Korp klassificeras som formellt subjekt inte att ifrågasättas. En poäng med att inte göra en kanske mer kvalitetssäkrad manuell genomgång av tusentals meningar, oaktat tiden det tar, är att man kan använda sig av taggningarna och samköra dem mot annoteringar om till exempel modersmål och betygsnivå och på så sätt lättare hitta intressanta samband mellan dessa parametrar.

Det är rimligt att utgå från hypotesen att helfrasinläring av expletivt *det är* med efterföljande predikativ eller egentligt subjekt är den mest frekventa i början av inlärarens utveckling och att sedan repertoaren utvidgas med andra verb. En

jämförelsevis låg förekomst av ordföljdsfel borde även det tyda på en form av helfrastillägnande.

Thyberg (2020: 80) redovisar i sin avhandling resultatet av sin analys av sammanlagt 9000 meningar ur svenskspråkiga textkällor av olika slag, vilket visar att lite drygt hälften av alla *det* är icke-referentiella. Det framkommer också ur hennes undersökning att det finns grännsfall av referentialitet utgörande mellan en och två procent.

Thyberg (2020: 70) bjuder även på ett test av icke-referentialitet: Ett icke-referentiellt *det* kan inte betonas, till exempel, och inte heller kan det bytas ut mot ett demonstrativt pronomen (jfr. Josefsson 2009: 145). Man kan också prova att byta ut *det* mot *allt* eller *det hela*. Nedan illustrerar det andra exemplet att *det* inte kan vara både betonat (illustrerat med kursiv) och expletivt.

(20) Jag hörde att du tränar mycket. *Det* är bra för hjärtat.

(21) \**Det* är bra för hjärtat med träning.

Vad som klassas som icke-referentiellt är dock inte trivialt. Thyberg (2020) för en diskussion om hur *det* utan tydlig antecedent kan referera antingen till ett textsammanhang eller till en ”mindre avgränsad situation” utanför texten. I dessa fall fungerar betoningstestet dåligt, men dessa breda referenser kommer räknas till de referentiella *det* och sorteras därför bort från de meningar som Thyberg senare analyserar i egenskap av icke-referentialitet.

## 6. Material

Studien utgår från ett källmaterial bestående av antal korpusar av L2-texter innehållande inlärtartexter. De har beteckningarna *SW1203 v1*, *TISUS v1*, *SpIn v1*, samt *SweLL v1 original* och *SweLL v1 target* (hädanefter refererade till utan ”v1”). De tre förstnämnda samlas som subkorpusar under beteckningen *SweLL-pilot* (en beteckning som fortsättningsvis inte kommer refereras till). De två senare samlas under beteckningen *SweLL-gold*. En viktig del av *SweLL-gold* korpusen (en

beteckning som fortsättningsvis inte kommer refereras till) är att den existerar i en så kallad *original*- och *target*-form. *Original*-versionen innehåller texterna inskrivna och inlagda som de är med alla avvikelser, medan *target* innehåller samma texter men korrigerade för att hamna nära normativ svenska (Volodina & Mohammed 2021). Därmed säger den numeriska skillnaden mellan dessa något om avvikelserna i inläroartexterna jämfört med normalsvenska. Sökningar på *target*-korporusen används vid utförandet av uppsatsens felanalysdel. Antalet essäer är 502.

Dessutom används en korpus av texter skapade på gymnasiet språkintröduktion, *SpIn*, och texter skrivna på avancerad nivå, *TISUS* (Volodina et al 2016).

L2-texterna är i sin helhet skyddade men nås efter tillstånd via Korp-verktygets internetadress.

Eftersom syftet är att undersöka hur ED används vid olika språkfärdighetsnivåer och eventuellt också att ge tentativa svar på hur konstruktionen utvecklas i inläroarspråket drar jag nytta av att – beroende på textkorpus – söka på (i) förekomster hos texter representerande en specifik kunskapsnivå (nybörjare, fortsättning, avancerad eller CEFR). Det förekommer också (ii) texter skrivna av specifika deltagare vid tre olika tillfällen där den longitudinella utvecklingen skulle kunna studeras på individnivå. Den senare möjligheten har inte använts. Korpusen benämnd *SW1203* uppfyller båda kriterierna och innehåller 141 uppsatser skrivna av 52 studenter varav 35 skrivit alla de tre proven. (Volodina et al 2016).

Vid sökning på alla *det* som pronomen får man i *SweLL-original* 2698 träffar, och i *SweLL-target* 2907. Vid sökning på ”det” som formellt subjekt (vilket inte skiljer sig åt i sökträffshänseende från att söka bara på *formellt subjekt*) får man 1 006 respektive 1 169. Statistiken från Thyberg (2020: 80) på vanliga svenska, icke-inläroartexter visar på runt hälften referentiella och hälften icke-referentiella *det* (och de formella subjekten är ju icke-referentiella, men inte alla icke-referentiella subjekt är formella). Den högre andelen referentiella *det* kan kanske ha med nybörjarkaraktern hos *SweLL*-texterna att göra.

Det finns möjligen en viss nackdel med *SW1203*, allt annat lika, eftersom det är färre individer i *SW1203*-korporusen vilket gör att avvikelser på individnivå kan slå igenom mera än om det varit 141 uppsatser av 141 individer. Det finns en nackdel också för *SweLL-original* och det är att det är ont om texter skrivna av individer på

(riktig) nybörjarnivå, vilket gör det svårare att säga något säkert kring förekomsterna i dessa texter.

Som nämnts tidigare förekommer det även internationellt att studier på inläratexter bygger på texter skrivna av personer som producerat dessa i anslutning till högre studier (Gonzalves 2019: 7). Det är också fallet med Hammarberg (2008).

## 6.1. Genrer

Den använda texttypen kan påverka vilka förekomster av olika ED-konstruktioner som påträffas. Språklig konstruktion som aktivt val fyller en funktion på samma sätt som val av texttyp anpassat för syfte och mottagare. I texter som bygger på att inläraren ska producera någon form av omdöme i en argumenterande text är det rimligt att tänka sig en högre frekvens av adjektiv (jfr. Hammarberg 2008: 94)

Vid val av jämförelsekorpus med L1 svenska är det därför viktigt att komma så nära man kan. Till den delen av uppsatsen som behandlar användningen av ED hos inlärare jämfört med texter skrivna på svenska, har jag valt att använda en korpus av texter från *Bloggmix 2015* och från internetforumet *Familjelivs* underforum *samhälle*. Ett forum av detta slag har texter av en formalitetsgrad liknande essäerna som skrivits av andraspråkseleverna och innehåller som forumtext ett mått av beskrivning, resonemang och argumentation. Här kommer framför allt den rent kvantitativa förekomsten av olika konstruktioner av ED i respektive korpus(ar) utgöra underlag för vad som utmärker inlärarspråket på respektive nivå. Dessa öppna korpusar är tillgängliga via Korpvertygets internetadress.

Man kan även tänka sig att formaliteten är lägre i den här jämförelsekorpusen men det bedömer jag gäller företrädesvis ordval och är inget som torde påverka förekomsten av konstruktionen i sig.

Det har visat sig att det inte är genomförbart att liksom sökningar i *SweLL* hitta ED genom att söka på *formellt subjekt* och *det*. Sökresultatet blir mycket lägre jämfört med det förväntade. Det material som Thyberg (2020: 68, 80) utgår från består bland annat av korpusen *Bloggmix 2014* och *Bloggmix 2015* som tillsammans utgörs av lite drygt 1,2 miljoner token. I det färdiganalyserade slumpmässiga urvalet hon utförde från dessa korpusar visar det sig att pronominet *det* var referentiellt i 50,3 % av fallen

och icke-referentiellt i 48,1 %”. (I akademiska texter var det överskott på icke-referentiella *det*). Mot bakgrund av dessa siffror borde det bli 600 000 träffar vid sökning på *formellt subjekt* och *det*, givet 1 258 050 *det* som pronomen, men i själva verket fick jag 137 285 träffar, eller 10,9 %. I stället för egna sökningar hade jag därför resultatet från Thyberg (2020) som referens för den kvantitativa förekomsten av expletivt *det* i L1-texter.

En annan sak som har bäring på genrer, och vilket kan påverka jämförelsernas validitet, är att inlärare skriver inom olika genrer inom olika kurser. På mycket tidiga nivåer skriver de till exempel inte argumenterande texter, eftersom de inte har den språkliga kompetensen till det ännu, utan ofta mera beskrivande texter. Man skulle till exempel kunna tänka sig att i enkla beskrivande texter finns det mer av presenteringskonstruktioner med verbet ”finnas”, medan det i argumenterande texter, som ju uttrycker värderingar, har en nödvändig del av konstruktioner med postponerat led, där värderingen antingen ligger i predikativt attribut eller i egenskapen hos substantivet som egentligt subjekt. Men detta har betydelse bara om man speciellt vill undersöka den kvantitativa relationen mellan postponerat led och presentationskonstruktioner, vilket inte kommer ske i denna studie.

Ofta blir det relativt få förekomster i absoluta tal vilket gör att procenttalen, speciellt när de är så låga som ensiffriga, ska ses som approximativa.

## 7. Resultat med diskussion

### 7.1. Förekomster

#### 7.1.1. *Utbrytningskonstruktionen*

Vad gäller referentialitet för *det* i utbrytningskonstruktionen förefaller det som *det* inte kvalificerar sig som icke-referentiellt i det av Thyberg (2020:70) presenterade *testet* som refererades till i tidigare avsnitt, eftersom vi i exemplet nedan *både* kan betona *det* och byta ut *det* mot ett demonstrativt pronomen.



(22) det (här) är fängelset som ska anta nya säkerhetsrutiner.

I Thybergs (2020: 77) studie av de icke-referentiella *det*-konstruktionerna, räknar hon även in utbrytningsliknande presenteringskonstruktioner bland utbrytningskonstruktionerna; dessa kallar hon ”presenteringsutbrytningar”. Det är alltså inte möjligt att använda hennes siffror som jämförelse vid jämförelser av vanlig utbrytning L1-texter mot våra L2-korpusar. SAG (4: 62, 518) talar om presentering med relativsats som ”besläktad” med utbrytningskonstruktionen, men skiljer sig från den senare genom att den ger *ny information* i relativsatsen.

(23) Det är barnen som ska tömma diskmaskinen den här veckan.

Samma information men icke-fokuserad:

(24) Barnen ska tömma diskmaskinen den här veckan.

Vid sökning i Korp på förekomster av den här konstruktionen måste man vara öppen för att till exempel subjektet kan vara en längre nominalfras, även om det typiska för konstruktionen är att det är ett substantiv eller pronomen. Man kan också tänka sig att ett satsadverbial, mer eller mindre fritt placerat, kan figurera. För att det ska bli hanterbart var det inte möjligt att konstruera sökningarna som med säkerhet hittade *alla* varianter av utbrytningskonstruktioner.

Man kan, precis som med *SweLL*-korpusen söka direkt på formellt subjekt, men antalet träffar blir orimligt få i ljuset av vad Thybergs (2020) undersökning ger vid handen vad gäller proportionerna mellan referentiellt och icke-referentiellt *det* som pronomen. Vid sökningarna har därför *det* som pronomen används och en söksträng som korresponderar mot den gängse utbrytningskonstruktionen.

I Hammarbergs (2008: 93f.) studie av *det är*-frasen i *muntligt* inlärarspråk konstaterar han att utbrytning (som han benämner med ”satsklyvning”) är betydligt mindre vanlig bland inlärare jämfört med infödda: 2,0% mot 8,2% av alla *det är*-satser (eller ”fragment” som han kallar dem). Vad gäller andel av token (antal ord) handlar det om 0,045% för inlärare 0,15% för infödda. Inlärarna är i hans undersökning

”kvalificerade”, som han skriver; de var inskrivna på universitetets preparandkurs och ”motiverade”, och det longitudinella insamlandet sträckte sig från det då var nybörjare till en nivå, som motsvarar en eller två år innan den nivå som krävs för högskolestudier Hammarberg (2008: 82).

Efter genomgång av träffar och sorterat bort ett antal referentiella *det* från sökresultatet ger en sökning i våra korpusar vid handen 207 instanser av traditionell utbrytningskonstruktion.

Även om det är fullt möjligt med utbrytningskonstruktion i en och samma mening med samordnade led och bisatser, har jag valt att presentera frekvensen mot antalet meningar snarare än token, då jag tänker mig att det normalt finns en utbrytningskonstruktion per fullständig mening. Frekvensen av utbrytningskonstruktioner fördelar sig mellan korpusarna enligt följande med avseende på andel av samtliga meningar: I *SW1203* 1,17%, *SweLL*-original 0,82% och *SpIn* 0,35%. För *TISUS* finns inte något angivet antal meningar, bara token, men om man extrapolerar utifrån att *TISUS*-texterna är skrivna på avancerad nivå och har ett argumenterande innehåll och därmed något fler token per mening, skulle vi hamna på kanske 2–3%. Relaterat till token blir andelen *totalt* för alla korpusar 0,080%, vilket går lättare att jämföra med Hammarberg (2008) som just redovisar förekomster mot andel av token. En mer valid jämförelse med Hammarbergs ”kvalificerade” inlärare skulle vara *TISUS* med texter skrivna av inlärare på avancerad nivå; det utgör då 0,15%, vilket sammanfaller med frekvensen hos infödda talare. Min förmodan är att den relativt högre frekvensen kan ha att göra med de mera argumenterande texterna som produceras på den nivå som *TISUS*-korpusen, där fokusering kan komma till speciell nytta, men det kan också ligga i den skillnaden mellan skriftligt och muntligt när det kommer till användning.

Eftersom det inte är möjligt att direkt jämföra med Thybergs (2020) resultat om frekvens, och inte heller med Hammarbergs (2008) *muntliga* resultat, måste resultatet från det här tyvärr begränsas till iakttagelser om den relativa skillnaden mellan de olika korpusarna. Man kan dock konstatera att mönstret med ökad språklig kompetens korresponderar med ökad användning av utbrytningskonstruktionen.

### 7.1.2. *Expletivt det med predikativt attribut och med-fras.*

I denna konstruktion står alltså *det* som subjekt och ett (ofta bedömande) predikativ. Det som omtalas och bedöms är rektionen i prepositionsfras med *med*. Den kan skrivas om med rektionen som subjekt med och utan ED.

(25) Det är besvärligt med cyklister på trottoarerna.

(26) Cyklister på trottoarerna är besvärligt.

Konstruktionen förekommer inte alls i *SpIn*, eller *SW1203*, och två gånger i hela *SweLL-original*-materialet. För att vara säker på att inte missa något fall gjordes sökningen på alla *det* som pronomen. I *TISUS* förekommer den *en* gång i en allt annat än typisk version där rektanten utgörs av ett ensamt *det*:

(27) [Ä]ven om man arbetade kanske lite för mycket var det inget fel med det.

(*TISUS*)

Thybergs (2020: 90) analys av sammanlagt 3000 meningar med icke-referentiellt *det* ur ett urval av korpusar resulterar i en fördelning bestående av endast 1,8% icke-referentiella *det* med *med-fras*-konstruktioner. Att konstruktionen inte heller existerar i sin typiska form (med verbet *vara*) med mer än *ett* atypiskt exempel i vårt *SweLL*-material är därför väntat. Att värdera, fälla ett omdöme, vilket skrivuppgifterna i våra korpusar i stor utsträckning inbjuder till, görs uppenbarligen genom andra sätt än med ED och *med-fras*.

Ur Hammarbergs (2008) analys av frasen *det är* i *inlärartal* kan vi tyvärr inte hämta några kvantitativa uppgifter om *med*-konstruktionen. ED skulle hamna under det som Hammarberg klassificerar som ett framåtppekande *det*, en egenskap det delar med de formella subjekten i allmänhet.

### 7.1.3. *Expletivt det med presentering och med postponerat led*

Sökning i Korp på *det* som formellt subjekt och egentligt subjekt gav träffar även på *det* vi klassificerar som postponerat led eftersom *det* postponerade leDET – bisats eller infinitivfras – har rollen som ett egentligt subjekt. Snarare än titta på hur inlärare hanterar ED i sig självt, var tanken att vi ska undersöka hur ED används i sammanhanget av ett egentligt subjekt.

Däremot är expletivt *det* i den betydelse vi hämtat från SAG, mer än formellt subjekt. Det jag valt att intressera mig för är dock ED som formellt subjekt inte som, till exempel, opersonligt subjekt vid utsagor om väder och allmänna tillstånd.

Det utmärkande för presenteringskonstruktionen är att den innehåller både ett formellt och ett egentligt subjekt, och som beteckningen antyder används konstruktionen för att introducera ny information.

- (28) Numera finns det olika levnadsformer man kan välja mellan utan att en ses som bättre än en annan. (*SW1203*)

Vid sökningar i Korp räknas, som nämnts, även hela led som egentligt subjekt, inte bara de nominalfraser som brukar tjäna som typexempel på egentligt subjekt i presentationskonstruktioner. I sökresultaten kom därför både presentationskonstruktioner och konstruktioner med postponerat led att ingå. Formellt subjekt korresponderar med egentligt subjekt i den taggning som Korp används mot, så vi fann en nästan fullständig överensstämmelse mellan formellt subjekt och egentligt subjekt: 1 171 mot 1 143 i *SweLL-target*. En fördel med *target*-versionen, även om den är en korrigerad text, är att *det* är lättare att hitta de typiska konstruktionerna som annars inte skulle bli upptäckta på grund av språkliga fel; man kan ju inte söka på *det* som inte finns om felet man söker handlar om något som är uteblivet.

Här kommer inte frekvensen av presentering och med postponerat led i L2-texterna jämföras med L1-texter, utan dessa båda konstruktioner behandlas i stället under felanalysen nedan.

#### 7.1.4. Expletivt ”det” och passiv konstruktion

I SAG talar man i sammanhanget av expletivt *det* och passiv om ”underförstådd mänsklig agens” (SAG 4: 57). I korpusen *SweLL*-original är det få verb som har en tydlig mänsklig agens, snarare är det verb angränsande till opersonlig betydelse. Det återfinns ett drygt tiotal passiva verb, bland dem *saknas*, *sägs*, *betraktas* och *skapas*. De utgör ungefär en procent av de formella subjekten med *det*. I den bloggtextkorpus Thyberg (2020: 91) använder sig av, utgör de passiva konstruktionerna 2,4%. Vad gäller olika färdighetsnivåer ger sig konstruktionen tillkänna uteslutande i texter på avancerad nivå, inte på fortsättnings- eller nybörjarnivå. Detta mönster bekräftades av sökningar i *SW1203* ned totalt *tre* förekomster på B1–C1 nivå, och att *SpIn* med texter representerade på A1–B2-nivå inte innehåller några passiva konstruktioner med ED överhuvudtaget. I *TISUS*, som bara innehåller texter på avancerad nivå, finns 14 exempel. Det motsvarar 3,0% av de formella subjekten i *TISUS*.

Det fanns alltså god överensstämmelse frekvensmässigt mellan L1-texter och de mera avancerade inlärtexterna representerade i *TISUS*.

## 7.2. Felanalys: Expletivt *det* som formellt subjekt och fokusfel utbrytningskonstruktionen

### 7.2.1. Expletivt *det* som formellt subjekt

På 70-talet etablerades i svensk andraspråkforskning begreppet ”platshållartvång” som benämning på företeelsen att alla svenska satser kräver ett subjekt och ett finit verb som en konsekvens av att ordningsföljden mellan dem är avgörande. Existensen av ett expletivt *det* kan sägas vara en konsekvens av detta, nämligen att det behövs ett *det* även när detta *det* inte refererar; på samma sätt skulle man kunna argumentera för kravet på finit verb: det behövs även om det bara är för att förbinda subjektet med predikativet som innehåller information (Källström 2012: 147–150).

För språk som ligger nära svenskan är det vanligt med subjektstvång, det subjektstvång som leder till icke-referentiella expletiva *det*. Men för språk som

arabiska, persiska, kurdiska – eller det mer närliggande spanska – finns det inget subjektstvång. I jämförelse med de femton jämförelsespråken som utgör vanliga inlärarespråk i Sverige och behandlas i Källström (2012), är det bara svenskan som har obligatoriska förekomster av det icke-referentiella, expletiva *det* (Källström 2012: 151–153). Man måste därför se det som en markerad struktur.

Som nämnts ovan existerar *SweLL*-korporuserna i en originalversion – *SweLL original* – där texten transkriberats som den är, och i en så kallad *target*-version – *SweLL-target* där avvikelserna i originaltexten har korrigerats till en godtagbar svenska. Skillnaden i storlek mellan de båda versionerna är 147 842 mot 151 851, vilket ger en nettodifferens på 4009 till den större *target*-versionens fördel. Det kan förstås vara såväl tillägg som uteslutning av token men man får anta att det framför allt har fattats token som skapat detta nettoöverskott i den korrigerade textversionen.

För att få en uppfattning om hur det ser ut i Korp och med annoteringsverktyget Svala med vilket man gjort ”normaliseringar” (ordvalet från Wiren et al. 2009) och satt ut koder för respektive fel, infogas följande figur:

Figur 1. Från Svala – ”normaliseringar” åskådliggjorda grafiskt



I den här felanalysen har jag studerat frekvensen av *uteblivet subjekt* i fallet där ett *det* som formellt subjekt, alltså ED, behövs för idiomatisk svenska enligt kriterierna för normaliseringen, eftersom en lägre frekvens för dessa uteblivna subjekt jämfört med andra subjekt skulle kunna tyda på att ED i de här fallen har etablerats i inlärares

språk som helfras, det som med engelskspråkig terminologi oftast går under beteckningen ”chunk” (Saville-Troike & Barto 2017: 80f.). En högre frekvens av uteblivna subjekt kan tyda på att konstruktionen är svårare för inläraren att hantera på korrekt sätt än vad andra subjekt är.

De kategorier av subjekt som man kan söka på i Korp är *formellt subjekt*, *egentligt subjekt*, *övriga subjekt* (*subjekt med infinitiv* ger träffar på infinitiven så det alternativet saknar relevans.) ED som formellt subjekt och egentligt subjekt förutsätter varandra och har direkt att göra med de ED-konstruktioner uppsatsen behandlar.

*Övriga subjekt* är helt enkelt subjekt som inte faller under de två förra specialfallen. Av de här subjekten är det 305 subjekt som ursprungligen saknades bland de totalt 17 609 vanliga subjekt som finns i den korrigerade textversionen *target*, vilket utgör andel av 1,7%. Antalet ”formella subjekt” i form av ED är i *target* 1 169 av vilka 171 ursprungligen saknades, vilket utgör en andel av 7,4 %. Nedbrutet till de olika kunskapsnivåerna och typen av subjekt blir resultatet följande:

Tabell 1. Procent, uteblivet formellt subjekt

Nivå / subjektstyp	uteblivna expletivt <i>det</i> antal / procent	uteblivna övriga subjekt antal / procent
<i>Nyborjare</i>	48 / 11,9%	115 / 2,1%
<i>Fortsättning</i>	12 / 9,8%	56 / 2,5%
<i>Avancerad</i>	27 / 4,1%	134 / 1,3%
<i>Totalt</i>	171 / 7,4%	305 / 1,7%

Ibland har det skett en mer omfattande justering av syntaxen i *target*, vilken har markerats med annan kod än den för saknat subjekt (se figur 1 ovan), även om justeringen innefattar (bland annat) just en konstruktion med tillägg av uteblivet formellt subjekt. Så man kan säga att de saknade formella subjekt som kvantitativt redogörs för här är, är sådana där det redan finns en ansats till en konstruktion med ED.

Det verkar alltså som om satskonstruktioner som involverar ED är svårare att hantera för inlärare än satskonstruktioner med vanligt subjekt, i alla fall när det gäller den del av hanteringen som handlar om svenskans subjektstvång.

### 7.2.2. *Expletivt det i presentering och postponerat led och inlärares modersmål*

Här har undersökts frekvensen av samma konstruktionsfel som ovan, nämligen uteblivet formellt subjekt, ett fel som korrigerats i target-versionen av *SweLL*, men med jämförelse mot *modersmål* – eller möjligen språkkunskaper eftersom flera språk förekommer under annoteringen ”modersmål” – och mot *graden av progression* som yttrar sig i en lägre frekvens av dessa fel på ökad ungefärlig kunskapsnivå.

Tyvärr finns det i materialet få inläratexter på *fortsättningsnivå*. Språken har valts ut efter hur välrepresenterade de är för att få ett så pålitligt utfall som möjligt. Persiska och dari har slagits ihop eftersom det finns en utpräglad korskompetens i de båda språken. Texter skrivna av personer som endast kan dari ingår också bland dessa. Arabiska har tagits med också i kombinationer med andra språk, till exempel arabiska och engelska och arabiska och kurdiska. Detsamma gäller ryska och spanska.

Tabell 2. Procent, uteblivet formellt subjekt

Nivå / Språk	Arabiska +	Persiska/Dari +	Spanska +	Ryska +
<i>Nyborjare</i>	13,5%	18,6%	18,8%	11,1%
<i>Fortsättning</i>	11,1%	<12,5%*	—**	—**
<i>Avancerad</i>	5,1%	4,9%	4,2%	6,5%

\*) Inget belägg    \*\*) inte representerad på fortsättningsnivå

Ett språk som hade hög representation på fortsättningsnivå var tigriska, vilket gav en felprocent på 8,3%. Tyvärr var de absoluta talen mycket låga på de övriga kunskapsnivåerna så det är svårt att sätta in detta data i ett större sammanhang.

Samtliga språk har det gemensamt att de inte har ett subjektstvång så som svenskan har, men i övrigt är de inte närbesläktade (Källström 2012: 151ff.). Det hade varit önskvärt att jämföra med ett språk ur den germanska språkfamiljen men de absoluta talen för träffarna var alltför låga. Möjligen är det någon indikation att vid sökning på



alla inlärartexter med tyska som författarspråk och lemgrammet *finnas* i *TISUS*, en korpus som bara innehåller texter på avancerad nivå, förekommer endast *ett* uteblivet *det* av 48 förekomster av *finnas*.

- (29) Åt andra sidan finns forskning som tyder på att det är mycket viktigt att har ett liv fylld med meningsfullt arbete. (*TISUS*)

Vad vi kan se av resultatet vad gäller modersmål är att det sker en förväntad progression i form av utförda fel, men speciellt värt att notera att det inte skiljer nämnvärt mellan språken, trots avstånden mellan dem. Men det kan också sägas vara ett förväntat med tanke på vad som nämnts i anslutning till felanalys i teoriavnittet: påverkan från modersmål när det gäller syntaktiska fenomen är inte stort (jämfört med till exempel fonologiska och lexikaliska företeelser).

### 7.2.3. *Utbyteskonstruktionen – fokusfel*

Jag har behövt gå igenom träffarna manuellt eftersom det inte går att göra sökningar i Korpus som utöver träffar på utbyteskonstruktioner inte också får med ett slags presentationskonstruktioner med *referentiellt det* och *är* och *som* som bisatsinledare till satsattribut, till exempel:

- (30) Det är en form av stress som vi kan definitivt undvika (*TISUS*)

Med avstamp i en iakttagelse av Ekerot (1979: 80f.; 2011: 105) om att inlärare kan behärska en utbyteskonstruktion syntaktiskt, men samtidigt används den på fel sätt, ska vi studera lite närmare de framgallrade utbyteskonstruktionerna. Till skillnad från presenteringskonstruktionerna, även de med utbytesliknande struktur (de som Thyberg [2020] räknar in i utbyteskonstruktionen men i övrigt bortser från i sin studie), är utbyteskonstruktionens syfte att fokusera redan känd information. Jag kan inte – bortsett från att konstruktionen kan ha andra grammatiska avvikelser – hitta mer än *ett* tveksamt exempel på felfokusering som ger sig tillkänna på meningsnivå.

- (31) Det tar 30 min till stan vi satt i bussen och det var en man som sitter bakom oss. (*SweLL original*)

Här blir det tydligt att informationen om *en man* introduceras i historien samtidigt som den fokuseras på det sätt som utbrytningskonstruktionen gör. Fortsättningen är intressant för samma konstruktion återkommer senare i berättelsen, men då med ett referentiellt *det*.

- (32) plötsligt vi hörde någon som skräck med högt röst och säger stoppp ! Alla vändade sig för att källa vem det var. Det var mänen som satt backom oss och bussöfren stanade bussen (*SweLL-original*)

Som tidigare nämnts om *tema* och *rema*, är analysen av informationsaspekten av fokuseringen något som på ett mera heltäckande sätt behöver göras utanför sats- och meningsnivå. Så de följande iakttagelserna och resonemangen bygger på meningsnivå, utifrån de träffar som erhållits genom korpsökningar.

Min genomgång av korpusarna *SW1203*, *SweLL-original*, *TISUS* och *SpIn* visade att denna fokuserande funktion verkar bemästras i stor utsträckning när den används.

### 7.3. Fördjupning: *det är* och *det finns* som högfrekventa helfraser

För att närma mig frågan om helfrasinlärning och förekomsten av ”chunks” i det *skriftliga* inlärarspråket utifrån frekvens och exponeringen av konstruktionerna i L1 svenska har jag valt att begränsa mig till fraserna *det är* och *det finns*. Till att börja med kan man konstatera att dessa två intar en särställning vad gäller frekvens jämfört med andra möjliga intransitiva verb för konstruktion med egentligt subjekt.

Tabell 3. Antal och andel *det är*, *det finns* av expletivt *det* + verb

Fragment/Korpus	Bloggmix 2015	Akademiska texter: sällskapsvetenskap	Familjeliv: samhälle
<i>D/det är</i>	10 347	3142	466 840

<i>D/det finns</i>	7790	4015	303 264
<i>Totalt: är, finns</i>	18 137 / 60,6%	7157 / 63,7%	770 104 / 63,3%
<i>Totalt:</i>	29 946	11 231	1 216 835

Vid sökning på korpusen *Bloggmix 2015* utgörs 60,6% av alla formella *det* + verbkonstruktioner av antingen *det är* eller ”*det finns*”, med en 32,8% övervikt för *det är* över *det finns*. I korpusen *Akademiska texter: samhällsvetenskap* är motsvarande siffror 63,7%, men med 27,8% övervikt för *det finns* istället. Med sociala medier och korpusen *Familjeliv: Samhälle* och 63,3% för *det är* och *det finns* tillsammans, är vi tillbaka till en övervikt för ”*det är*” igen med på 53,9%.

I sig själva visar dessa andelar inne mera än att de verkligen är satsfragment som inläraren mycket ofta möter i text, och möjligen visar det också att texten när den tjänstgör som åsiktsmedium premierar *det är*-konstruktioner över *det finns*-konstruktioner. Genreaspekten verkar slå igenom ganska tydligt.

Finns det några spår i inlärartexterna av att dessa ytterst vanliga satsfragment kommer in i texterna som ett slags odelbara enheter, snarare än som satsled uppbyggda av två delar? Vad som spelar roll här, som hypotes, är också vid vilken nivå dessa eventuellt fyller en funktion, och möjligen också utbildningsbakgrund och på hur med vilken strategi inläraren tar sig an och reflekterar om ett främmande språk. En delfråga till detta kunde vara om *det är* i någon fas av inlärningsprogressionen *det* uppträder mer frekvent.

Medan de svenska referenstexterna för L1 ovan rör sig kring 60–64% *det är*- och *det finns*-fragment, lägger sig samtliga L2-texter som vi använder oss av sammantaget mellan 68–75%. Dessa högre siffror behöver inte säga annat än en helt förutsägbar konsekvens av att inlärare har en mindre repertoar av uttryck med ED. Jämförelsen mellan texterna i korpusarna är då kanske av lite större intresse.

Tabell 4. Antal och andel *det är*, *det finns* av expletivt *det* + verb

<i>Fragment/Korpus</i>	SweLL-original	SW1203	TISUS	SpIn
<i>D/det är</i>	281 / 37,0%	185 / 43,5%	110 / 36,8%	27 / 21,8%

<i>D/det finns</i>	270 / 35,5%	132 / 31,1%	94 / 31,4%	63 / 50,8%
<i>Totalt: är, finns</i>	551 / 72,5%	317 / 74,6%	204 / 68,2%	90 / 72,6%
Totalt: 760	760	425	299	124

*SpIn* har jämfört med övriga L2-korpusar den största andelen inlärartexter på nybörjarnivå, och är den enda med texter på A1/A2. Men textmängden här är liten jämfört med andra nivåer bland L2-korpusarna. Till skillnad från övriga tre korpusar har *SpIn* fler *det finns* än *det är*, dessutom med ordentlig marginal, 133%.

Redan resultaten omkring uteblivet ED i tidigare avsnitt har pekat i riktning mot att ett *det* allt annat än självklart slår följe med *finns*. Jag har valt att titta på *det finns*-fragmentet i nybörjarkorpusen *SpIn*. Jag ville se om det fanns spår av ett mer ”mekaniskt” användande av *det finns*, likt det som etablerats hos tidigare inlärare av till exempel Axelsson (1989).

Det visade sig att konstruktioner utan *det* är vanliga i *SpIn*: Av totalt 119 förekomster av *F/finns*, på samtliga kunskapsnivåer, står 49 utan *det*, vilket motsvarar 41,1%. Om det är helfraser som lagras i minnet och som även används vid skriftlig produktion, skulle man rimligen inte finna alls många uteblivna subjekt. I alla fall kan vi konstatera att *det finns* inte används så i skrift. Av ett slumpmässigt urval av 200 meningar ur *SpIn* med *är* är det endast ett exempel på att *är* saknar *det* som subjekt, expletivt eller inte:

- (33) nu jag är inte nervös och nu vet jag att här är [det] inte som min himland.  
(*SpIn*)

Bland samma meningar är det samtidigt 39 *det är*. På muntligt inlärmaterial är vad gäller *det är*, som vi sett tidigare, mycket av terrängen kartlagd av Axelsson (1989) och Hammarberg (2008); detta formelartade fragment ses där som en positiv resurs i talsituationen, vilket trots sin relativa grammatiska komplexitet i kombination med nominalfraser, predikativa attribut och existentialkonstruktioner, används med lätthet (Hammarberg 2008: 103f.).

Det finns alltså en stor skillnad mellan hur *det är* hänger ihop jämfört med *det finns* trots att de båda är nästan lika vanliga i texter som inlärare möter, beroende på genre.

Eftersom denna följsamhet kan iakttas så tidigt som nybörjartexterna i *SpIn* finns det ingen anledning att tro att ”det” skulle tappas bort i högre utsträckning på högre inlärnivåer.

Det förefaller alltså som att *det är* uppträder som ett kluster likt i talspråk.

## 8. Avslutning

### 8.1. Slutsats

De svar på forskningsfrågorna som med stöd av den återgivna teoretiska och ramen svagt avtecknar sig med via sökningarna i inlärningskorpusarna, pekar i riktning mot att tillägnandet av konstruktioner med expletivt *det* inte verkar följa en radikalt annan bana än annat jämförbart språktillägnande av L2.

Det verkar vara så att expletiva *det* har saknats i högre utsträckning än andra subjekt i inlärtexterna, vilket kan tyda på att konstruktioner med expletivt *det* är svårare att behärska, vilket på felanalysens grund bekräftar att det är en markerad struktur givet inlärningsförstaspråk.

Expletivt *det* i olika konstruktioner börjar fylla sin funktion från början i kunskapsutvecklingen med en växande repertoar med ökad progression, och en jämnt avtagande frekvens av avvikelser. Ovanliga ED-konstruktioner i L1 är också ovanliga eller obefintliga i L2.

*Tema* och *rema* har nämnts i uppsatsen som en viktig förklaringsbakgrund till konstruktionen med expletivt *det*, men har i övrigt inte följts upp. En analys av detta skulle kräva att man studerar större stycken av text än meningar. Den grundläggande överensstämmelsen i frekvens som nämnts ovan, visar på att inlärningsförstaspråket har haft samma behov av informationsförskjutning i den fullständiga satsen som förstaspråksanvändare.

Kvantitativt avspeglas en större andel av de två vanligaste satsfragmenten *det är* och *det finns* i förhållande till andra möjliga *det* och verbkombinationer jämfört med svenska texter vilket inte torde ha någon annan orsak än tillgången till ett större ordförråd att konstruera satser med ED med.

Resultaten från studien tyder på att den speciella ställning som *det är* som ett slags enhet vilket etableras tidigt i *mundligt* inlärarespråk, som påvisats i flera studier, kan spåras även i *texter*. Med *speciell* menas här att det inte kan hänföras till ren frekvens, då – i alla fall i *texter* – *det finns* (med eller utan *det*) ibland är lika vanlig som *det är*, med viss variation för genre. Detta kan säga något om frågan om relationen mellan skriftligt och muntligt, vilket berördes i teoridelen, nämligen att inte bara den muntliga produktionen kan användas som adekvat mått på kompetensnivå, utan även den skriftliga produktionen.

De på senare år allt mer elaborerade teorierna om andraspråkstillägnade, vilka behandlades i teoridelen med avseende på input och output, implicit och explicit kunskap samt förhållandet mellan skriftlig produktion och muntlig produktion, skulle egentligen behövt större material och ett mer omfattande studium än den föreliggande uppsatsen. Men den högfrekventa förekomsten av ED-konstruktionen bär ändå på potentialen till ett rimlighetsresonemang, nämligen att behärskning framförallt kommer från implicit kunskap genom en rik input ända från språkutvecklingens tidiga fas. I den mån det har kunnat göras jämförelse med muntligt material är det inget som tyder på att den naturliga tillgången till explicit kunskap vid textskrivande har påverkat frekvens eller avvikelser i en linjär progression, avvikelser som till exempel en överanvändning av en inlärd regel skulle kunna skapa.

## 8.2. Vidare forskning

Subjektsanvändning i bredare bemärkelse i svenskan som inlärarespråk än den som studerats i den här uppsatsen är av gammalt datum. Korpusstudier kan genom sin kvantitativa fördel föra den vidare. För detta är det viktigt med ett korpusmaterial som ständigt utvidgas kvantitativt över många språk, både språk närstående till svenskan och de som typologiskt står långt ifrån. En naturlig fortsättning på det som utförts i denna studie är att med huvudsakligen samma frågeställningar i botten, och med hjälp av korpusar rikta uppmärksamheten mot enskilda språk- och språkgrupper och annan information som redan finns annoterat i anslutning till texterna.

Utanför korpus-studier tänker jag mig att med kvalitativt inriktade studier, att expletivt *det* kunde studeras med fokus på inläring och tillägnande av språk med

metoder som överlappar området mellan lingvistik och pedagogik. Det kunde till exempel vara möjligt finna svar på *hur*- och *varför*-frågor om inlärares användning av ED genom observation eller intervjuer i lärandesituationer.

## 9. Litteratur

- Abrahamsson, Niclas (2009). *Andraspråksinläring*. Lund: Studentlitteratur.
- Axelsson, Monica (1989). Helfraser och ramar – ett viktigt inslag på andraspråksinlärares väg mot ett kreativt regelsystem. I: Hyltenstam, Kenneth & Inger Lindberg (red.). *Första symposiet om svenska som andraspråk. Vol. 1.: Föredrag om språk, språkinläring och interaktion*. Stockholms universitet, Centrum för tvåspråkighetsforskning. S. 191–202.
- Ballardini, Kerstin, Stjärnlöf, Sune & Viberg, Åke (1986). *Mål: svenska som främmande språk. Svensk grammatik : svensk grammatik på svenska*. Stockholm: Natur och kultur.
- Doughty, C. J. (2003). Instructed SLA: Constraints, compensation, and enhancement. I: Doughty, C.J. & Long, M.H (eds.). *The handbook of second language acquisition*. Oxford: Blackwell. S. 256–310.
- Ekerot, Lars-Johan (1979). Syntax och informationsstruktur. I: Hyltenstam, Kenneth (red.). *Svenska i invandrarperspektiv*. Lund: Liber. S. 79–108.
- Ekerot, Lars-Johan (2011). *Ordföljd, tempus, bestämdhet: Föreläsningar om svenska som andraspråk*, 2:a uppl. Malmö: Gleerups.
- Ellis, Nick C. & Wulff, Stefanie (2020). Usage-based approaches to L2 acquisition. I: VanPatten, Bill & Williams, Jessica (eds.) *Theories in second language acquisition: an introduction*, 2nd ed. S. 63–82.
- Ellis, Rod (2008). *The study of second language acquisition*. 2. edn. Oxford: Oxford university press.
- Engdahl, Elisabet (2010). Vad händer med subjektstvånget? Om det-inledda satser utan subjekt. *Språk & stil. Tidskrift för svensk språkforskning*, 20. S. 81–104.
- Fasth, Cecilia & Kannermark, Anita (1996). *Form i fokus: övningsbok i svensk grammatik Del A*. Lund: Folkuniversitetets förlag
- Fasth, Cecilia & Kannermark, Anita (1997). *Form i fokus: övningsbok i svensk grammatik Del B*. Lund: Folkuniversitetets förlag
- Fasth, Cecilia & Kannermark, Anita (1998). *Form i fokus: övningar i svensk grammatik. Del C*. Lund: Folkuniversitetets förlag.



- Forsskåhl, Mona (2009). *Konstruktioner i interaktion: de e som resurs i samtal*. Helsingfors: Svenska litteratursällskapet i Finland.
- Gonzalves, Lisa (2019). *Understanding and Supporting Emergent Writing in L2 Adults Developing First-Time Print Literacy*. University of California. (dr-avh.)
- Hammarberg, Björn (2008). Konstruktioner som produkt och process – en studie av hur L1- och L2-talare utnyttjar «det är». *Nordand: Nordisk tidskrift för andrespråksforskning*, 1 (3). S. 79–107.
- Josefsson, Gunlög (2009). *Svensk universitetsgrammatik för nybörjare*, 2:a uppl. Lund: Studentlitteratur.
- Källström, Roger (2012). *Svenska i kontrast*. Lund: Studentlitteratur.
- Levy Scherrer, Paula & Lindemalm, Karl (2014). *Rivstart A1 + A2 Textbok*. 2. upplagan. Stockholm: Natur & Kultur.
- Levy Scherrer, Paula & Lindemalm, Karl (2015). *Rivstart B1 + B2 Textbok*. 2. uppl. Stockholm: Natur & kultur.
- Lindström, Jan (2008). *Tur och ordning. Introduktion till svensk samtalsgrammatik*. Stockholm: Norstedts Akademiska Förlag.
- Makaela, Leketi (2004). Differential error types in second-language students' written and spoken texts: implications for instruction in writing. *Written communication*, 22 (4). S. 368–385.
- Melin, Lars & Lange, Sven (2000). *Att analysera text: Stilanalys med exempel*. Lund: Studentlitteratur.
- Polio, Charlene (2012a). The acquisition of second language writing. I: Gauss, S. M., and Mackey, A (eds). *The Routledge handbook of second language acquisition*. New York: Routledge. S. 319–334.
- Polio, Charlene (2012b). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21. S. 375–389
- SAG = Teleman, Ulf, Staffan Hellberg & Erik Andersson (1999). *Svenska Akademiens grammatik*. Vol. 1–4. Stockholm: Norstedts ordbok.
- Saville-Troike, Muriel & Barto, Karen (2017). *Introducing second language acquisition*. 3rd ed. Cambridge: Cambridge university press.

- SO = Svenska akademien and Lexikaliska institutet (2009). *Svensk ordbok*: utgiven av Svenska Akademien. Stockholm: Nordstedt distributör. <https://svenska.se>
- Sundman, Marketta (1980). *Existentialkonstruktionen i svenskan*. Åbo: Meddelanden från Stiftelsens för Åbo akademi forskningsinstitut.
- Marketta, Sundman (1987). *Subjektval och diates i svenskan*. Diss. Åbo: Åbo Akad.
- Thyberg, Kajsa (2020). *Det-konstruktioner i bruk: En systemisk-funktionell analys av satser med icke-referentiellt det i modern svenska*. Göteborg: Göteborgsstudier i nordisk språkvetenskap 41.
- Volodina, Elena, Mohammed, Samir Ali (2021). *Korp searches in L2 material (SweLL): SweLL hands-on exercise/manual*.
- Volodina, Elena, Pilán, Ildikó, Enström, Ingegerd, Llozhi, Lorena, Lundkvist, Peter, Sundberg, Gunlög, Sandell Monica (2016). *SweLL on the rise: Swedish Learner Language corpus for European Reference Level studies*. Proceedings of LREC 2016, Slovenia.
- Williams, Jessica (2012). Classroom research. I: Gauss, S. M., and Mackey, A (eds). *The Routledge handbook of second language acquisition*. New York: Routledge. S. 541–554.
- Winlund, Anna (2021). Writing practices of recently immigrated adolescent emergent writers: A study from a language introductory school in Sweden. *Journal of Second Language Writing*, 51. S. 1–12.
- Wirén Mats, Arild Matsson, Dan Rosén, Elena Volodina. (2019). *SVALA: Annotation of Second-Language Learner Text Based on Mostly Automatic Alignment of Parallel Corpora*. CLARIN-2018 post-conference volume. LiUP Press. <http://www.ep.liu.se/ecp/159/023/ecp18159023.pdf>

Korpusverktyg

Korp <http://spraakbanken.gu.se/korp/> (2022-12-27)