



MEDICINSKA FAKULTETEN

Lunds universitet

Avdelningen för logopedi, foniatri och audiologi

Institutionen för kliniska vetenskaper, Lund

Hörfrekventa innehållsord i svenska bilderböcker - förskolebarns förmåga att producera vanligt förekommande innehållsord

Annie Karlsson & Anna Svensson

Logopedutbildningen, 2019

Vetenskapligt arbete, 30 högskolepoäng

Handledare: Ketty Andersson & Kristina Hansson

Sammanfattning

Syfte: Syftet var att bidra till insikter om den lexikal-semantiska utvecklingen genom att identifiera vilka innehållsord samt semantiska domäner barn förväntas behärska i förskoleåldern. **Metod:** En korpus av populära bilderböcker för svensktalande barn sammanställdes för att undersöka vilka ord som är högfrekventa samt vilka semantiska domäner dessa ord tillhör. För att ta reda på i vilken utsträckning orden är förvärvade utformades ett benämningstest där barnen skulle mobilisera innehållsord i meningskontext. Testet utfördes på nio barn med typisk språkutveckling. **Resultat:** Antalet identifierade högfrekventa innehållsord var 93, av vilka majoriteten utgjordes av verb. Vidare identifierades totalt 27 domäner, varav de största utgjordes av *läge och rörelse, natur* och *omdöme*. Avseende testresultaten var verb svårast att benämna medan substantiv och adjektiv erhöll fler korrekta svar hos deltagarna. **Slutsatser:** De mest frekventa orden i korpusen var verb, vilka hade en generell betydelse. Barn i detta åldersspann förefaller ha tillägnat sig mer än hälften av de 50 undersökta orden i hög utsträckning och resterande ord i viss omfattning. Det gick inte att påvisa något tydligt samband mellan ett ords frekvens i korpusen och hur många barn som kunde benämna det. Studiens tillvägagångssätt och resultat kan bidra till större förståelse för barns lexikal-semantiska utveckling.

Söktermer: barnspråk, korpus, lexikal-semantisk utveckling, semantiska domäner.

Abstract

Purpose: The purpose was to increase the knowledge about children's lexical-semantic development by identifying which content words and semantic domains they are expected to master in the preschool age. **Method:** We compiled a corpus of popular children's books for Swedish speaking children to investigate which words occur most frequently and which semantic domains these words belong to. We designed a naming test in which the children were to mobilize content words in a context. The test was performed on nine children with typical language development. **Results:** The most frequent words consisted of verbs. In total, 27 domains were identified, of which the largest consisted of *space and motion*, *nature* and *opinion*. Furthermore, the verbs appeared to be the most difficult to name. **Conclusions:** The most common content words in the corpus consisted of verbs with a general meaning. Children in this age span appeared to have acquired more than half of the 50 assessed words to a large extent and the remaining words to some extent. No evident relationship was found between a word's frequency and the number of children naming it. The approach and results of this study can contribute to a greater understanding of children's lexical-semantic development.

Key words: children language, corpus, lexical-semantic development, semantic domains.

INNEHÅLLSFÖRTECKNING

INLEDNING	1
BAKGRUND	1
DEFINITIONER	1
TEORIER OM LEXIKAL SEMANTIK	2
FÖRUTSÄTTNINGAR FÖR LEXIKAL UTVECKLING	3
DE FÖRSTA ORDEN	4
PRODUKTION KONTRA FÖRSTÅELSE	4
UTVECKLING AV SEMANTISKA DOMÄNER	4
FAKTORER SOM PÅVERKAR LEXIKAL UTVECKLING	5
ELICITERINGSMETODER FÖR LEXIKAL FÖRMÅGA	5
LÄSNINGENS BETYDELSE FÖR ORDINLÄRNING	5
SYFTE	6
FRÅGESTÄLLNINGAR	7
METOD	7
MATERIAL FÖR KORPUSEN	7
PROCEDUR FÖR KORPUSEN	7
DATAANALYS	8
MATERIAL FÖR PILOTSTUDIEN	8
DISTRIBUTION ÖVER BÖCKERNA	8
LÅGFREKVENTA ORD	9
UTFORMNING AV TESTET	9
METOD FÖR PILOTSTUDIEN	9
TESTDELTAGARE OCH REKRYTERING	9
TESTFÖRFARANDE	9
FORSKNINGSETISKA ÖVERVÄGANDEN	10
RESULTAT	10
DE HÖGFREKVENTA INNEHÅLLSORDEN I KORPUSEN	10
SEMANTISKA DOMÄNER	12
UTVALDA ORD TILL BENÄMNINGSTESTET OCH RESULTAT FRÅN PILOTTESTNINGEN	13
GEMENSAMT FÖR HELA GRUPPEN	17
KARAKTÄRISTISKA DRAG FÖR RESPEKTIVE ÅLDERSGRUPP	17
FREKVENNS	17
DOMÄNER	17
ORDKLASSER	18
SAMMANFATTNING AV RESULTATET	19
DISKUSSION	19

FREKVENSORDLISTA.....	19
ORDKLASSER.....	19
DOMÄNER.	19
PILOTTESTNING	20
KARAKTÄRISTISKA DRAG FÖR RESPEKTIVE ÅLDERSGRUPP.....	20
GEMENSAMT FÖR HELA GRUPPEN.....	21
FREKVENSENS BETYDELSE.....	22
ETABLERADE DOMÄNER.....	22
VERBENS KOMPLEXITET.....	22
METODDISKUSSION.....	23
ELICITERINGSSTRATEGIER.....	23
FÖRSLAG PÅ KOMPLETTERANDE TEST.....	23
TIDSÅTGÅNG.....	24
BEGRÄNSNINGAR.....	24
SLUTSATSER.....	24
IMPLIKATIONER FÖR LOGOPEDISKT ARBETE OCH FRAMTIDA FORSKNING	25
TACK	25
REFERENSER.....	26

BILAGA 1 - BÖCKER

BILAGA 2 - HÖGFREKVENTA INNEHÅLLSORDS FREKVENNS OCH DISTRIBUTION ÖVER BÖCKERNA

BILAGA 3 - BENÄMNINGSTEST

BILAGA 4 - BILDSTÖD

BILAGA 5 - INFORMATIONSBREV TILL VÅRDNADSHAVARE

BILAGA 6 - MEDGIVANDEBLANKETT

Inledning

Barn tillägnar sig sina första ord omkring ett års ålder. Det är i huvudsak med hjälp av ord vi förmedlar tankar. I takt med att ordförrådet expanderar behöver det så småningom omorganiseras för att barnet ska kunna hantera det ökande antalet ord och skapa struktur (Nettelblatt, 2007). Förmågan att strukturera ords betydelser i domäner (semantiska fält) anses vara en viktig förutsättning till att barn lär sig språk effektivt (Gärdenfors, 2014). Ord som är betydelsemässigt relaterade kan definieras som ett semantiskt fält. För att ge ett exempel faller orden *koka*, *baka*, *friter* och *rosta* inom begreppet *matlagning*, som utgör ett semantiskt fält (Lehrer, 1985). Genom samtal med vuxna får barnet information om ordets form, betydelse, uppbyggnad och användning (Clark, 2009). Vuxna benämner föremål och aktiviteter samt förklarar hur orden inom en kategori relaterar till varandra (Clark & Wong, 2002). En annan väsentlig källa för språkutvecklingen är gemensam läsning. Genom böcker exponeras barnet för fler unika ordtyper än vad barnriktat tal för sig självt erbjuder. Det ger barnet gynnsamma möjligheter till att lära sig ord och deras betydelser (Montag, Jones & Smith, 2015).

Bakgrund

För att få insikt i förskolebarns lexikal-semantiska utveckling kommer det i föreliggande studie att skapas en korpus av bilderböcker, avsedda för barn i åldrarna tre till sex år. Utifrån korpusen ska sedan en frekvensordlista sammanställas över de mest frekventa innehållsorden. Med utgångspunkt i dessa ord kommer därefter deras tillhörande semantiska domäner att identifieras. Vidare kommer ett antal av de högfrekventa innehållsorden att ingå i en pilotstudie, där några förskolebarn kommer testas för att se i vilken utsträckning de behärskar orden. I följande avsnitt redogörs grundläggande begrepp inom lexikal-semantik samt utvecklingen av denna förmåga.

Definitioner

Barn förvärvar successivt ord och bygger därmed upp sitt *lexikon*. Lexikonet fortsätter att förändras och utvecklas hela livet i takt med att vi möter nya företeelser i vår omgivning (Nettelblatt, 2007). Varje ord innehar en betydelse, vilken kan beskrivas med termen *semantik*. Semantik berör den språkliga betydelsen hos såväl ord som morfem, fraser och satser. *Lexikal semantik* kallas studiet av enskilda ords betydelser och betydelserelationer mellan ord (Nationalencyklopedin, u.å).

Ett ord kan förekomma i flera olika former, till exempel ordet *stol* kan böjas som *stol*, *stols*, *stolen*, *stolens*, *stolar*, *stolars*, *stolarna*, *stolarnas*. Den här uppsättningen ord kan härledas tillbaka till ett *lemma* (i detta fall *stol*). Många lemman har mer än en betydelse. För att kunna urskilja en betydelse hos ett lemma används enheten *lexem*. Lexem kan definieras som grundbetydelse hos ett ord (Järborg, 2007). Vidare kan det vara värt att förklara termen *lexikaliserad fras*, som innebär att orden i frasen står i, mer eller mindre, fast förbindelse med

varandra. Ordförbindelsens betydelse är således enhetlig och kan inte härledas från de enskilda orden (Johansson-Kokkinakis, 2007).

Lexem kan ytterligare delas in i flertalet ordklasser, som i sin tur kan placeras i en av två huvudkategorier: *innehållsord* respektive *funktionsord*. Det är innehållsorden som förmedlar den huvudsakliga betydelsen hos ett yttrande medan funktionsorden talar om vilka relationer innehållsorden har till varandra inom ett yttrande. Till innehållsord räknas ordklasserna substantiv, verb, adjektiv, och till viss del adverb. Till funktionsorden räknas bland annat pronomen, artiklar, prepositioner och konjunktioner. Ytterligare termer som kan behöva förklaras är löpord (eng. *tokens*) som är det totala antalet ord och unika ordtyper (eng. *types*) vilket innebär antalet olika ord (Nationalencyklopedin, u.å).

Teorier om lexikal semantik

Kognitiv semantik. Det finns olika teorier kring vad ordbetydelse egentligen är. Denna uppsats baseras på teorin *kognitiv semantik*. Enligt den kognitiva traditionen är betydelser mentala representationer (Gärdenfors, 2014), det vill säga betydelser finns i människors tankar (Paradis & Willners, 2011). Det läggs, mer specifikt, betoning på relationerna mellan kommunikativa handlingar (däribland språkuttryck) och individens mentala föreställningar av deras betydelser. Formandet av dessa mentala representationer sker inte oberoende av den yttre omgivningen, utan i samspel med den. Betydelse skapas således i kommunikation och uppkommer i interaktion mellan samtalspartners (Gärdenfors, 2014).

Med utgångspunkt i det kognitiva synsättet, hävdar en del forskare att lexikala enheter fungerar som ledtrådar, snarare än att de "har" betydelser. Dessa används för att göra slutledningar för att därefter skapa förståelse och mening (Cruse, 2002). Beroende på i vilka konstruktioner språkliga enheter används, i både talat och skrivet språk, aktiveras särskilda konceptuella strukturer. Lexikal betydelse är relationen mellan lexikala enheter och det sammanhang de används i, det vill säga *när* de används. Vidare är lexikala betydelser föränderliga och begränsas av sin universella och konventionella betydelse samt kontext (Cruse, 2002; Paradis, 2003). I denna mening, får således ord sina specifika betydelser i kommunikation samt utifrån de konstruktioner och kontexter som de används i (Paradis & Willners, 2011).

En specifik teori inom kognitiv semantik är Gärdenfors teori om semantiska domäner. Ett centralt begrepp inom teorin är *begreppsmässiga rymder*, 'Conceptual Spaces', med vilka han menar att vi organiserar information i våra tankar till ett format som kan beskrivas i geometriska eller topologiska termer (Gärdenfors, 2014). Mera specifikt består begreppsmässiga rymder av dimensioner vilka representerar grundläggande egenskaper (kvaliteter) hos föremål. Eftersom olika föremål besitter en viss kvalitet i olika utsträckning kan de kartläggas till olika punkter i en rymd, likt ett koordinatsystem. På basis av kvalitetsdimensionerna kan stimuli tolkas som lika eller olika. Exempelvis kan vi bedöma toner utifrån deras tonhöjd vilket möjliggör för oss att rangordna tonerna efter likhet. Avstånden mellan representationerna av till exempel föremål är ett mått på hur lika de är varandra vad gäller en specifik egenskap, till exempel färg, vikt eller form (Gärdenfors, 2018). Således gör användningen av geometri och topologi det möjligt att beskriva närhet och avstånd i en rymd. "Om en punkt x är närmare punkt y än punkt z, så är x mer lik y än z" (De Geer & Gunnarsdottir, 2013, s. 3). Vidare bildar vissa dimensioner kluster, vad Gärdenfors

kallar *domäner*. Inom exempelvis domänen *färg* finns de olika dimensionerna *nyans*, *mättnad* och *ljusstyrka* (Gärdenfors, 2014). En domän representerar sålunda en särskild uppsättning egenskaper (Gärdenfors, 2018).

Enligt Gärdenfors (2018) innebär den semantiska utvecklingen att barnet reducerar och organiserar mängder av komplex sensorisk information till mer hanterbara strukturer. Detta åstadkoms genom att barnet identifierar olika sorters invarianter, det vill säga konstanta eller oföränderliga egenskaper hos företeelser. Det leder i sin tur till skapandet av primära, perceptuella domäner. Efter att de primära domänerna konstruerats detekterar vår hjärna kovarianser för olika särdrag (dimensioner), det vill säga identifierar föränderliga, samvarierande egenskaper. Det här görs för att kunna skapa begrepp av fenomen i vår omgivning. Skapandet av begrepp gör det möjligt att kategorisera, kommunicera och resonera.

Gärdenfors (2014) hävdar vidare att vår begreppsmässiga kunskap är organiserad i domäner och att detta är en förutsättning för att kunna utveckla språk. Att tillägna sig en ny domän är mer kognitivt krävande än att tillägna sig nya begrepp i en redan befintlig domän. I sin redogörelse för hur semantiska domäner utvecklas uppger Gärdenfors att detta framför allt sker genom att domänerna separeras från varandra och att nya bildas genom en redan befintlig sådan, snarare än att nya domäner adderas. Ett exempel utgörs av Piagets (1972) studier, i vilka barn skulle välja mellan två glas med läsk, det ena var smalt och högt medan det andra var brett och lågt. Han fann att barn under fem års ålder tenderade att välja det glas som var högst trots att det breda glaset innehöll mer läsk. Det här beror på att barn i den här åldern inte kan separera volymen av en vätska från dess höjd. Det är först senare de lär sig att vätskevolym inte alltid korrelerar med höjd. Barnen har då lärt sig att separera dimensionen *höjd* från dimensionen *volym* och därmed tillägnat sig en ny domän.

Strukturalistisk semantisk teori. En annan teori som är relevant för den valda testmetoden i föreliggande studie är *strukturalistisk semantisk teori*. Det strukturalistiska synsättet anser att språk är ett autonomt inomspråkligt system bestående av ordrelationer. Teorin skiljer mellan paradigmatiska och syntagmatiska relationer (Paradis, 2012). Med paradigmatiska relation menas relationen mellan ord som befinner sig inom samma betydelsekategori (Nettelbladt, 2007). Det inkluderar synonymer (till exempel *kallt* och *kyligt*), antonymer (till exempel *varmt* och *kallt*) och hyponymer (till exempel *katt* och *djur*). Syntagmatiska relationer syftar däremot på linjära relationer mellan ord inom meningar. En syntagmatisk relation innebär då att orden förekommer tillsammans med andra ord i ett yttrande och genom detta skapas associationer (Cruse, 1986). Syntagmatisk relation är av särskild relevans vid undersökning av ordproduktion i meningskontext, som i föreliggande studie (Paradis, 2012).

Förutsättningar för lexikal utveckling

Barnet behöver skapa representationer för sådant som de ser, hör, känner och smakar för att kunna registrera fenomenet när det återkommer. Förmågan att lagra representationer i minnet är en förutsättning för att kunna kategorisera eller organisera erfarenheter (Clark, 2009). För att kunna föreställa sig inre bilder krävs det i sin tur att andra kognitiva förmågor är etablerade. Det leder därefter fram till den så kallade benämningsinsikten. Det är först när barnet tillägnat sig denna som det förstår att ord är symboler för något (Nettelbladt, 2007).

De första orden

Tidpunkten för när barn producerar sina första ord varierar något, vanligast är omkring 12 till 15 månaders ålder. De första orden är avhängigt av sådant som sker omkring barnet, till exempel människor som de träffar dagligen samt återkommande aktiviteter och rutiner. Detta gör att ordförrådet kan variera betydligt från individ till individ. För barn som växer upp i samma kultur går det dock att finna vissa likheter i ordförrådet (Håkansson, 2014).

Människor förvärvar ord i en meningskontext och vi utnyttjar det språkliga sammanhanget för att härleda ordbetydelser (Bloom, 2000). Ordinläring i kontext sker bäst genom att vi hör ett ord användas i en konversation (Nagy & Herman, 1987). Troligtvis lärs många ord in på detta sätt. Fördelarna är att konversationen oftast sker i en icke-språklig kontext (utöver det språkliga i kommunikationen), att talaren till viss del är medveten om lyssnarens förkunskap, och att lyssnaren har möjlighet att ställa frågor (Bloom, 2000).

Barn är kapabla att lära sig ord efter att de endast hört ordet en gång och gör sällan stora misstag, till exempel att de skulle tro att *mjolk* betyder *gaffel* (trots att den vuxna tittar på gaffeln medan hen säger mjölk). Ord kan således läras in utan att barnet samtidigt ser dess referent (Bloom, 2000). Barn behöver heller ingen återkoppling för att lära sig ordbetydelser. Även om det är vanligt i västvärlden att föräldrar korrigerar sina barns tal, är det inte universellt (Lieven, 1994). Vidare benämner vuxna ofta föremål, till exempel *det här är en kaka*, men vanligtvis inte handlingar. När vuxna i barnets omgivning använder verb utförs oftast inte den avsedda handlingen samtidigt (Gleitman, 1990; Tomasello, 1992). Exempelvis, när vuxna använder ordet *öppna* gör de oftast det utan att öppna något, och tvärtom, när vuxna öppnar något benämner de sällan samtidigt vad de gör (Bloom, 2000).

Produktion kontra förståelse

Det är skillnad i antalet ord som barnet producerar respektive förstår. Det tar längre tid att lagra representationer för produktion jämfört med förståelse (Clark, 1993). Att producera ett ord kräver att barnet dels hämtar representationen för ordet och dels relevant artikulatorisk information (Clark, 2009). I en undersökning testade Goldin-Meadow, Seligman och Gelman (1976) tvååringars förmåga att förstå och producera 80 substantiv respektive 30 verb. Resultatet visade att de förstod 61 procent medan de kunde producera 37 procent. Skillnaden var större för verb jämfört med substantiv. Vidare kunde de endast producera ord som de även förstod (Clark, 2009). En studie som utfördes av Barrett (1995) visade att vid omkring 17 månaders ålder producerar barnet 45 ord medan det förstår 189 ord. Diskrepansen minskar ju äldre man blir (Nettelblatt, 2007) men skillnaden kvarstår upp i vuxen ålder (Clark, 2009).

Utveckling av semantiska domäner

Barns utveckling av semantiska domäner sker genom utvidgning och underindelning. Genom att reorganisera och ordna begrepp på överordnade, sidoordnade och underordnade nivåer kan barnet utveckla sitt vokabulär. Barn kan vid två års ålder dela in ord i sidoordnad och underordnad nivå och vid omkring tre års ålder använda samtliga nivåer (Nettelblatt, 2007). Barns tidiga ordinläring tar sin början i ett litet antal semantiska domäner, vilka är nästintill samma som de var för 50 till 100 år sedan. Barn pratar, precis som vuxna, om människor, föremål och aktiviteter som sker runt omkring dem (Clark, 1993). Andra semantiska kategorier som de 50 första orden, i flera språk, tillhör är djur (till exempel *hund*,

fågel, katt), mat/dryck (till exempel *banan, juice, äpple*), kroppsdelar (till exempel *öga, näsa, öra*), klädesplagg (till exempel *sko, strumpa, mössa*), fordon (till exempel *bil*), leksaker (till exempel *boll, bok*), hushållsföremål (till exempel *flaska*), rutiner (*hej, hejdå, nej, godnatt, bada, tack*) och aktiviteter såsom ljud effekter, riktning och tillstånd (till exempel *uh-oh, woof, moo, vroom, upp, ner*) (Fenson et al. 1994). Detta bekräftades i en studie av Luniewska et al. (2015) som undersökte tilläggnandet av 299 ord i 25 språk. Studien visade att barn till stor del förvärvar de undersökta orden i åldrarna tre till fem år och 95 procent av orden inföll i ovannämnda semantiska kategorier. När barnet har 50 till 100 ord i sitt expressiva ordförråd finns vanligen endast ett eller två ord i respektive semantiskt fält (Clark, 1993).

Faktorer som påverkar lexikal utveckling

Det är vedertaget att ju frekventare ett ord är desto tidigare lär sig barn det. Goodman, Dale och Li (2008) påvisade emellertid det motsatta, att ju mer frekvent ett ord är desto senare förvärvas det. Ett språk har vanligtvis fler unika substantiv än verb och endast en liten uppsättning av funktionsord. Ett funktionsord används därför mer frekvent jämfört med det genomsnittliga verbet, som i sin tur är mer frekvent än det genomsnittliga substantivet. Trots detta är substantiv en stor del av det tidiga lexikonet. Inom varje ordklass stämmer dock antagandet att högfrekventa ord förvärvas tidigare än lågfrekventa ord.

I en studie som gjordes av Hansen (2017) undersöktes vilken betydelse frekvensen har för ordinlärning. Frekvensen var den faktor, av ordklass, frekvens, avbildbarhet och ord med liknande fonologisk uppbyggnad, som spelade störst roll i barnriktat språk för vid vilken ålder barn förvärvar ett ord. I enlighet med detta menar Leonard (1998) att högfrekventa ord lärs in snabbare än lågfrekventa. Vidare är det lättare att tillägna sig ord som har liknande fonologisk struktur som de ord som redan finns i barnets lexikon. Ord som är lättare att skapa en mental representation av förvärvas även tidigare än ord som är svårare att föreställa sig. Detta kan vara en förklaring till varför substantiv dominerar över verb i barns tidiga lexikala utveckling (Hansen, 2017).

Eliciteringsmetoder för lexikal förmåga

För att kunna bedöma lexikal förmåga, och andra språkliga aspekter, hos barn används olika eliciteringsstrategier. När det gäller bedömning av lexikon är den vanligaste eliciteringsmetoden konfrontationsbenämning, vilken innebär att testledaren ber barnet att benämna bilder eller föremål. En annan eliciteringsmetod som kan användas för att bedöma lexikal förmåga är meningsifyllnad. Den här strategin innebär att testledaren påbörjar en mening som barnet sedan ska avsluta genom att fylla i det ord som fattas. Frågestrategi är ett annat exempel på eliciteringsstrategi. Testledaren ställer i det här fallet en fråga som barnet ska besvara med ett bestämt ord (Nettelbladt, 2007). I den här studien tillämpas eliciteringsstrategierna meningsifyllnad och frågestrategi.

Läsningens betydelse för ordinlärning

Barn lär sig språk på basis av den språkliga input de får från sin omgivning. Flera studier, däribland den utförd av Weisleder och Fernald (2014), visar att tal som riktas till barnet, snarare än tal mellan vuxna, är den allra viktigaste källan för den tidiga

språkinläringen. De ovannämnda författarna fann att tal från vårdnadshavare både hade en direkt och en indirekt påverkan på den lexikala utvecklingen. Genom att barnet blev mer exponerat för barnriktat tal gav det barnet fler möjligheter att lära sig nya ord. Det föreföll vidare ha en gynnsam effekt på barnets framväxande lexikala bearbetningsfärdigheter, som i sin tur gav vinning åt ordförrådstillväxten. Därtill har studier som undersökt individuella skillnader i barnriktat tal kunnat visa att den lexikala och syntaktiska variationen i vårdnadshavares tal predicerar utfallet av barnets språkinläring, däribland ordförrådet. Således ger mer variation bättre utfall (Hoff & Naigles, 2002; Rowe, 2012; Weizman & Snow, 2001).

En viktig källa till lexikal variation i input till barn är enligt Montag et al. (2015) bilderböcker, som kan definieras som böcker i vilka texten regelbundet kompletteras med bilder (Ganidou, 2007). Jämfört med barnriktat tal av samma mängd, innehöll bilderböckerna fler unika ordtyper (eng. *types*). Författarna konkluderade utifrån detta att bilderböcker kan vara en viktig källa för ordförrådstillväxten (Montag et al., 2015). Gemensam bokläsning har också visat sig ha positiv inverkan på den tidiga språkutvecklingen. Barn som läser fler böcker tillsammans med sina vårdnadshavare uppvisar bättre språkförmågor och även större framgång med läsinläringen jämfört med barn som exponeras för färre böcker (Bus, van Ijzendoorn & Pellegrini, 1995; Farrant & Zubrick, 2012; Payne, Whitehurst & Angell, 1994).

Läskunniga personer lär sig majoriteten av ord genom text, även om den skriftliga kontexten är mindre effektiv än konversationer. En elev som läser relativt lite (såsom elever som inte läser på fritiden) exponeras uppskattningsvis för en halv miljon ord per år genom läsning, varav 10 000 är okända för eleven (Nagy & Herman, 1987). I föreliggande studie förväntas de äldre barnen ha påbörjat sin läsutveckling och därmed exponerats för ord i såväl tal som text till skillnad från de yngre barnen som deltog i pilotstudien. Miller (1996) menar att läsning är det enda som kan förklara hur en del vuxna tillägnat sig väldigt rika ordförråd. Med tanke på läsningens betydelse för ordförrådsutvecklingen, är ett sätt att få en uppfattning om vilka ord som finns i barns lexikon att studera populära böcker inom ett angivet åldersspann. I nuläget saknas dock en sådan sammanställning av språkdata, en korpus, för svensktalande barn.

Syfte

I denna uppsats kommer en korpus att skapas av svenskspråkiga bilderböcker för barn i åldrarna 3 till 6 år. Anledningen till att korpusen baseras på böcker är, som nämnts, att bokläsning har en stor betydelse för ordinläringen. Utifrån korpusen ska en frekvensordlista sammanställas av de mest högfrekventa innehållsorden och dess tillhörande semantiska domäner. Korpusen ska sedan utgöra utgångspunkt för att fastslå i vilken utsträckning de högfrekventa orden är förvärvade i detta åldersspann. Uppsatsens syfte är att bidra till insikter om den lexikal-semantiska utvecklingen och hjälpa oss att identifiera vilka innehållsord samt semantiska domäner som barn förväntas behärska i förskoleåldern. Dessa insikter kan vara en utgångspunkt för bedömning av lexikal-semantisk förmåga hos barn samt för val av målord som ligger inom barnets proximala utvecklingszon vid intervention för barn med språkstörning.

Frågeställningar

- Vilka är de mest högfrekventa innehållsorden i populära bilderböcker för åldrarna tre till sex år?
- Vilka semantiska domäner tillhör dessa ord?
- I vilken utsträckning är orden förvärvade hos barn i förskoleåldern (3;9 till 6;1 år)?

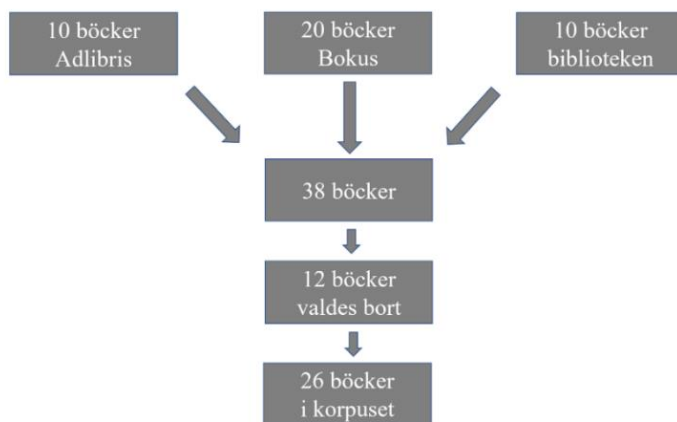
Metod

Material för korpusen

Underlaget bestod av 26 bilderböcker (bilaga 1) som valdes ut från en lista av de mest utlånade böckerna från sex bibliotek i Skåne samt bästsäljande listor från två av Sveriges största bokhandlare (Adlibris och Bokus). Adlibris topplista fanns tillgänglig via deras hemsida medan topplistorna från Bokus och biblioteken erhöles via mail. Topplistan från biblioteken var sammanslagen från samtliga folkbibliotek i nordöstra Skåne.

Procedur för korpusen

Med utgångspunkt från bokhandlarnas och bibliotekens topplistor införskaffades böcker till korpusen. En översiktlig bild för skapandet av korpusen ges i figur 1.



Figur 1. Flödesschema för skapandet av korpusen

Målet var att samla in cirka 30 böcker. Med hänsyn tagen till att samtliga böcker inte fanns tillgängliga bestod listan från början av 40 böcker för att säkerställa urvalsstorleken. Det ska nämnas att två av böckerna från bokhandlarnas topplistor överlappade. Listorna bestod således av 38 olika böcker. Av dessa böcker fanns två inte tillgängliga, varvid påföljande bok i topplistan valdes ut. Vidare exkluderades även tio av de tillgängliga böckerna på grund av att de var ljudböcker, endast innehöll bilder, bestod av fler än 250 sidor eller var svårskannade. Totalt samlades således in 26 böcker till korpusen, varav 18 kom från bokhandlarnas listor (åtta respektive tio vardera) och tio från bibliotekens lista. Samtliga böcker var avsedda för åldersgruppen tre till sex år och tillhandahölls genom bibliotek.

Efter att böckerna var insamlade skannade vi in texterna med hjälp av en skanner med funktionen optical character recognition (OCR) och c-pen reader pen scanner. Böcker som innehöll lite text skrevs in manuellt. Den text från böckerna som inkluderades i korpusen var den som var avsedd att läsas för barnet. Verser och refränger från visor som stod utskrivna räknades även med. Vi exkluderade egennamn, innehållsförteckning, förord, tack från författaren, instruktioner till läsaren, text i bilder och baksidestext. Texten från samtliga böcker överfördes därefter till datorprogrammet AntConc (Anthony, 2018).

Dataanalys. Med hjälp av AntConc (Anthony, 2018) skapade vi en sammanslagen frekvensordlista för samtliga böcker. Genom programmet rangordnade vi orden efter frekvens och alfabetisk ordning. En anledning till att orden ordnades alfabetiskt var för att identifiera ord som inte var tillräckligt frekventa i endast en böjningsform och således inte fanns med bland de mest högfrekventa orden. Ordet *öga* var till exempel inte frekvent nog för att vara högt upp i frekvensordlistan i AntConc men vid sammanföring av ordets samtliga böjningsformer (*öga, ögat, ögats, ögon, ögonen, ögonens*) blev dess förekomst högre. Det underlättade även lemmatisering (sammanföring av böjningsformer av ett ord), semantisk och syntaktisk disambiguering (semantisk tolkning och ordklassindelning), samt indelning i semantiska domäner. Mer specifikt, vid bearbetning av materialet stämde vi av att samtliga böjningsformer av respektive högfrekvent ord beräknades. Vi analyserade även ord som kunde ha flera betydelser för att ta reda på vilken som avsågs i sammanhanget, till exempel kan ordet *får* syfta på såväl ett verb som ett substantiv och ordet *få* kan vara en infinitivform av ett verb och ett adjektiv. Detta gjordes för att få en rättvisare bild av ordfrekvenserna (Järborg, 2007). Vi beräknade således frekvens utifrån såväl ordform som betydelse.

Utifrån listorna i AntConc (efter frekvens- och bokstavsordning) valde vi därefter ut de mest högfrekventa orden (93 ord) tillhörande ordklasserna substantiv, verb, adjektiv och adverb bildade ur adjektiv (bilaga 2). Vi bestämde att den totala förekomsten av respektive ord skulle vara minst 35 för att få inkluderas i uppsatsens frekvensordlista. Beslutet togs framförallt på grundval av att ett tillräckligt stort antal ord ville samlas in till den påföljande testningen. För att listan skulle bli jämförbar med orden i en ordbok blev det nödvändigt att sammanföra orden till deras grundform (lemma). Ord i ordböcker och ordlistor brukar stå i sin grundform, till exempel substantiv i obestämd singular och verb i infinitiv (Järborg, 2007).

Material för pilotstudien

Det primära syftet med uppsatsen var att framställa en korpus och en frekvensordlista. Uppsatsförfattarna ville även genomföra en pilotstudie för att undersöka huruvida barn i det angivna åldersspannet har förvärvat de ord som är mest frekventa i korpusen. Pilotstudien var således det sekundära syftet med uppsatsen. För att uppnå det syftet utformades och genomfördes ett benämningstest (bilaga 3).

Distribution över böckerna. Innan målorden till testet valdes ut beräknades samtliga högfrekventa innehållsords fördelning över böckerna (bilaga 2). Detta gjordes för att säkerställa att orden återfanns i flera av böckerna och minska risken för att någon bok skulle ha för stort genomslag. Vi ville försäkra oss om att ett ord förekom flera gånger i många böcker och att barn därigenom med större sannolikhet exponerats för det. För att få en överblick av distributionen av varje ord jämfördes förekomsten av orden i hela korpusen med förekomsten av orden i respektive bok. Därefter sattes gränsvärden för vilka ord från

frekvensordlistan som skulle användas i benämningstestet. Respektive ord skulle återfinnas i minst tio böcker och i minst åtta av dessa skulle det förekomma minst 2,5 procent av dess totala förekomst i böckerna. Gränsvärdet sattes till 2,5 procent då ordens relativa frekvens ofta befann sig omkring detta värde i respektive bok. Detta var därmed utslagsgivande för om ett ord skulle ingå i testet.

Den relativa frekvensen omvandlades sedan till absolut frekvens då denna kan skilja sig mycket åt beroende på hur stor den totala förekomsten för varje ord är. För att få en överblick av hur den absoluta frekvensen varierar för respektive ord angavs den lägsta och högsta förekomsten. Sammanlagt var högfrekventa ord med god spridning 65 i antal. Några av dessa inkluderades trots det inte i testet då de bedömdes vara svåra att illustrera och/eller elicitera (exempelvis *tycka, finnas*).

Lågfrekventa ord. I testet valde vi även att inkludera ord som var lågfrekventa (sju ord). Anledningen till detta var att kunna jämföra deltagarnas prestation mellan hög- och lågfrekventa ord. När vi valde ut dessa ord utgick vi från att orden skulle förekomma totalt fem gånger eller färre i korpusen, samt vara lätta att elicitera och illustrera. De utvalda lågfrekventa orden skulle dessutom ha en jämn fördelning mellan ordklasserna.

Utformning av testet. Testet konstruerades som en saga och skälen till detta var dels för att underlätta elicitering av målord och dels för att barn främst exponeras för ord i en meningssammanhang och inte isolerat. Innan målordet hade testats innehöll inte sagan ordet ifråga. För att förtydliga målorden kompletterades sagan med bilder (bilaga 4). Bilderna införskaffades via Darts kommunikation- och dataresurscenters (u.å) bildbank. Benämningstestet bestod slutligen av 50 målord, med en relativt jämn fördelning mellan ordklasserna.

Metod för pilotstudien

Testdeltagare och rekrytering. Deltagarna bestod av nio barn i åldrarna 3;9 till 6;1 år. En deltagare exkluderades från studien då hen var yngre än tre år. Deltagarna delades upp i två grupper utifrån ålder vid analys. Det gjordes på grundval av att gemensamma felsvar gick att identifiera inom de olika grupperna. Den yngre gruppen utgjordes av fem barn i åldrarna 3;9 till 4;4 år och den äldre gruppen av fyra barn i åldrarna 4;7 till 6;1 år. Sju av deltagarna var flickor och två var pojkar. Könstillfördelningen var relativt jämn inom grupperna. Samtliga hade typisk språkutveckling, var normalhörande och enspråkiga med svenska som modersmål. Denna information erhöles uppsatsförfattarna skriftligt från vårdnadshavarna. Deltagarna rekryterades genom en förskola. Vi kontaktade förskolechefen via mail och telefon i syfte att informera och få godkännande. Efter att godkännande erhållits informerades berörda pedagoger på förskolan. Dessa tillfrågade sedan vårdnadshavare och delade ut informationsbrev (bilaga 5) samt medgivandeblankett (bilaga 6).

Testförfarande. Testningen utfördes på förskolan, i ett avskilt rum. Testet administrerades genom att en av uppsatsförfattarna läste sagan för deltagaren samt visade bildstödet medan den andra uppsatsförfattaren registrerade deltagarens svar. För att elicitera målorden använde vi meningssifyllnad som eliciteringsstrategi i kombination med bildstöd. Bilderna visades en i taget på en surfplatta i samband med att deltagaren skulle benämna målordet. Om deltagaren inte förstod vilket målord vi syftade på (svarade felaktigt eller svarade inte alls) tillämpades frågestrategi. Tillämpning av frågestrategi skedde inte på ett

strukturerat, förutbestämt vis. Ifall rätt svar uppgavs genom denna eliciteringsmetod bedömdes det som korrekt. Felaktigt uttal eller böjningsform förbisågs. Inkorrekta svar noterades för att möjliggöra analys.

Forskningsetiska överväganden

Deltagarna tilldelades ett kodnamn för pseudonymisering. Kodnamnen förvarades i ett låst skåp på Avdelningen för logopedi, foniatry och audiologi och fanns endast tillgänglig för personer som var involverade i uppsatsen. Insamlade data från deltagarna bestod av blanketter. Såväl textmaterial från korpusen som testblanketter kommer destrueras efter det att uppsatsen är godkänd. Testningen påbörjades först när deltagaren hade en påskriften medgivandeblankett från båda vårdnadshavarna. I informationsbrevet framgick uppsatsens syfte och dess bakgrund samt testförfarandet. Vidare framgick det att medverkan var frivillig, pseudonymiserad och att det var tillåtet att avbryta deltagandet. Deltagandet medförde inga risker. En individuell nytta som uppsatsen kan ha för vårdnadshavare, och indirekt för deltagarna, är ökad insikt i barnets lexikal-semantiska utveckling. En nackdel med testningen kan vara att deltagarna upplevde situationen som kravfylld. Projektet har godkänts av den Etiska kommittén vid Avdelningen för logopedi, foniatry och audiologi, Institutionen för Kliniska Vetenskaper Lund, Lunds Universitet.

Resultat

I resultatet redovisas frekvens, ordklass och semantiska domäner för högfrekventa innehållsord samt några lågfrekventa ord i korpusen. Vidare presenteras de utvalda orden för benämningstestet utifrån deltagarnas svar och frekvens i korpusen. Totalt innehöll korpusen 54112 löpord (eng. *tokens*) och 7240 unika ordtyper (eng. *types*).

De högfrekventa innehållsorden i korpusen

I tabell 1 redovisas de 15 mest frekventa innehållsorden i korpusen. Orden presenteras i fallande frekvens.

Tabell 1. Högfrekventa innehållsord i korpusen, i fallande frekvens.

Ord	Frekvens	Ordklass
Vara	1557	Verb
Ha	623	Verb
Säga	570	Verb
Kunna	488	Verb
Ska	340	Verb

Få	332	Verb
Bli	304	Verb
Komma	304	Verb
Gå	269	Verb
Lite/liten	265	Adverb/adjektiv
Se	228	Verb
Göra	214	Verb
Ta	176	Verb
Mycket	173	Adverb
Stor	171	Adjektiv

De 15 mest frekventa innehållsorden bestod av tolv verb, två adverb och två adjektiv. Ett av orden kunde klassas som såväl adverb som adjektiv. Fördelningen av ordklass för samtliga högfrekventa innehållsord (93 ord) var 37 verb, 34 substantiv, 21 adjektiv och fem adverb. Sammantaget kunde fyra av orden klassas som både ett adjektiv och ett adverb. Orden med allra högst frekvens utgjordes av verb medan substantiv, adjektiv och adverb var mindre frekventa och således finns längre ner på listan. Det mest frekventa substantivet förekom först på plats 19 (ordet *dag*).

För att ta reda på vilka högfrekventa ord som är brett representerade i korpusen undersöktes deras fördelning över böckerna. I bilaga 2 presenteras distributionen för samtliga högfrekventa innehållsord. I den går det att utläsa att de 65 ord som var väldistribuerade, det vill säga hade en god spridning över böckerna, innefattade 33 verb, 13 substantiv, 18 adjektiv och fyra adverb. Tre ord kunde klassas som såväl ett adjektiv som ett adverb. Verb var således den vanligaste ordklassen bland dessa ord.

Ord som förekom mer än 175 gånger totalt i korpusen återfanns i nästintill alla böcker. Endast tre ord förekom i samtliga böcker (*vara*, *ha*, *säga*), vilka också var de tre mest frekventa. Orden med lägre frekvens hade tendens att förekomma i successivt färre böcker. Mönstret var däremot inte konsekvent. Exempelvis var orden *vatten* och *träd*, jämfört med andra substantiv, högfrekventa men förekom i relativt få böcker i förhållande till deras totala förekomst i korpusen.

Avseende antalet böcker där ordet förekom förhållandevis ofta (minst 2,5 procent) var siffrorna mer spridda och tenderade inte att sjunka ju lägre frekvens ordet hade. Det gick således inte att utläsa något tydligt mönster för hur dessa värden förhöll sig till varandra (till exempel att värdena gradvis sjönk) eftersom respektive värde är beroende av ordets totala förekomst.

Semantiska domäner

I tabell 2 presenteras de högfrekventa innehållsorden i sina respektive semantiska domäner. Blackwell (2005) har varit en inspirationskälla vid indelning av adjektiv och adverb medan Fellbaum (1998) har fungerat som utgångspunkt vid indelning av verb. Uppsatsförfattarna har själva delat in substantiven i lämpliga domäner i samråd med handledare.

Tabell 2. Indelning i semantiska domäner av samtliga högfrekventa innehållsord.

Domän	Ord inom domänen
Djur	Djur, fågel, hund, höna, katt, spindel
Fordon	Bil, cykel, tåg
Form & gestalt	Hel, rund
Färg	Grön, röd
Hushåll	Bo, hem, hus, säng
Kognition	Tro, tänka, veta, vilja
Kommunikation	Fråga, ropa, säga
Kroppsdelen	Huvud, mage, öga
Känsla	Glad, rädd, trött
Läge & rörelse	Börja, flyga, få, gå, hoppa, komma, köra, ligga, sitta, stå, ta, åka
Mat	Mat, äta
Mängd	Lite, mycket, många
Människa	Barn, gubbe
Natur	Blad, blomma, gräs, mark, skog, snö, träd, vatten, växt
Omdöme	Bra, gilla, kunna, måste, precis, riktig, tycka, veta
Perception	Se, titta
Socialt beteende & interaktion	Hjälpa, leka, snäll
Storlek	Lilla, liten, lång, små, stor

Svårighetsgrad	Lätt
Sömn	Somna, sova
Temperatur	Varm
Tid	Dag, kväll, natt, precis, vinter, vår
Tillhörighet	Ha
Tillstånd	Bli, finnas, göra, ha, ska, vara
Vikt	Lätt
Ålder	Gammal, ny, år
Övernaturligt	Spöke

Identifierade domäner var 27 i antal. Ett fåtal ord (*lite/liten, lätt, ha, precis, veta*) tillhörde två domäner eftersom de var semantiskt mångtydiga. *Lite/liten* kan syfta på såväl *mängd* som *storlek* och *lätt* kan tillhöra både domänen *svårighetsgrad* och *vikt*. Vidare kan ordet *ha* gälla såväl *tillhörighet* som *tillstånd*, medan *veta* kan avse både *kognition* och *omdöme*. Den domän som innehöll flest ord var *läge och rörelse* (tolv ord) och den näst största domänen var *natur* (nio ord). *Läge och rörelse* innefattade endast verb medan *natur* innehöll enbart substantiv. Det är värt att notera att majoriteten av domänerna innehöll ett till tre ord.

Utvalda ord till benämningstestet och resultat från pilottestningen

Vid utformning av benämningstestet togs dels hänsyn till att målorden skulle vara högfrekventa och dels att orden skulle vara väldistribuerade över böckerna. Målord med god spridning eftersträvades för att minska risken att någon bok skulle ha för stort genomslag. Av samtliga högfrekventa innehållsord (93 ord) var 65 väldistribuerade.

I benämningstestet inkluderades även sju lågfrekventa ord i syfte att jämföra deltagarnas prestation mellan hög- och lågfrekventa ord i korpusen. I tabell 3 presenteras dessa ord utifrån frekvens, i fallande led.

Tabell 3. Utvalda lågfrekventa innehållsord i korpusen.

Ord	Frekvens
Källare	5
Bädda	4
Moln	4

Slänga	3
Läskig	3
Korsord	2
Kräsen	1

De lågfrekventa innehållsorden innefattade två verb, tre substantiv och två adjektiv. De valdes ut med hänsyn till att frekvensvärdet inte fick överstiga fem, det skulle vara en jämn fördelning mellan ordklasserna samt vara lätta att illustrera och elicitera. De lågfrekventa orden tillhörde följande domäner: *förbrukning (slänga)*, *hushåll (bädda, källare)*, *omdöme (läskig)*, *socialt beteende och interaktion (kräsen, läskig)*, *syssla (korsord)* samt *väder (moln)*. Observera att ordet *läskig* kan tillhöra två olika domäner.

I benämningstestet exkluderades flera verb eftersom dessa var svåra att illustrera och elicitera. Benämningstestet bestod av 19 verb, 16 substantiv, 14 adjektiv och ett adverb. Fördelningen av ordklasser blev på så vis jämnare i benämningstestet, eftersom flera av verben exkluderades, jämfört med samtliga högfrekventa innehållsord. I benämningstestet avsågs betydelsen *mängd* gällande *lite/liten* och *vikt* avseende *lätt*.

I tabell 4 redovisas samtliga ord som inkluderades i benämningstestet. Orden presenteras efter antal och andel korrekt angivna svar av deltagarna. Det ges även exempel på felsvar från deltagarna.

Tabell 4. Målord i benämningstestet, antal/andel korrekta svar, frekvens i korpusen och exempel på deltagarnas felsvar.

Ord	Antal deltagare som svarade korrekt (procentandel)	Frekvens	Exempel på felsvar
Gå	9 (100)	269	
Lite/liten	9 (100)	265	Röd*
Stor	9 (100)	171	
Sitta	9 (100)	137	Väntar*
Många	9 (100)	122	
Träd	9 (100)	101	
Sova	9 (100)	97	
Flyga	9 (100)	81	Flaxar*

Hus	9 (100)	75	Brunt*, svart*
År	9 (100)	75	
Stå	9 (100)	74	Vara stor*
Rund	9 (100)	73	
Hjälpa	9 (100)	72	Dra upp*
Natt	9 (100)	72	Kväll*
Hoppa	9 (100)	71	
Ny	9 (100)	71	
Barn	9 (100)	66	Flicka*, pojke*
Öga	9 (100)	51	
Glad	9 (100)	48	
Huvud	9 (100)	40	
Varm	9 (100)	38	
Gräs	9 (100)	35	
Grön	9 (100)	35	Kladdig*, slemmig*
Moln	9 (100)	4	
Lång	8 (89)	135	Lågt*, stort, rakt*
Vatten	8 (89)	119	
Titta	8 (89)	84	Kikar, kollar*, spejar*
Ligga	8 (89)	80	Sover*
Hem	8 (89)	56	Tillbaks
Skog	7 (78)	69	Djungel*, träd
Källare	7 (78)	5	Trappa
Läskig	7 (78)	3	Mörkt, rädd
Få	6 (67)	332	Bli glad, ta, äta, öppna*

Se	6 (67)	228	Blinkar, kikar, tittar, spanar*
Mycket	6 (67)	173	Halva, hög, stor
Tänka	6 (67)	136	Gör, är*
Hel	6 (67)	128	Cirkel*, fyra*, rund*, stor
Gammal	6 (67)	69	En tant, mormor*, stor*
Ropa	6 (67)	69	Skriker*, säger
Kväll	6 (67)	39	Dag
Slänga	6 (67)	3	Lägga dom i lådan, lägger, kastar
Åka	5 (56)	81	Går, kör
Bädda	5 (56)	4	Byter lakan, gör det fint, städar
Dag	4 (44)	140	Ljust, morgon, solen
Fråga	4 (44)	53	Gissar*, pratar, säger, tänker*
Bo	3 (33)	50	Bygger, gubbe, går ut
Gilla	3 (33)	37	Räcker upp tummen, tycker om, älskar, är bra
Lätt	2 (22)	39	Inte tunga, liten, små, stora, väger inte mycket
Korsord	2 (22)	2	Spel, pussel, problem
Kräsen	0 (0)	1	Allergisk, full, gillar inte det, inte hungrig, stopp, tåler inte

* Målordet benämns alltid korrekt efter frågestrategi

Genomsnittspoäng samt lägsta och högsta testpoäng beräknades, både för hela gruppen och för respektive åldersgrupp. Dessa värden presenteras i tabell 5.

Tabell 5. Genomsnittspoäng samt lägsta och högsta testpoäng inom hela gruppen och varje åldersgrupp.

Grupp	Genomsnittspoäng	Lägsta och högsta poäng
Hela gruppen	40	33-46
Yngre gruppen	36	33-39
Äldre gruppen	45	43-46

Deltagaren med lägst testpoäng fick 33 medan den med högst fick 46 poäng. De äldre deltagarna (4;7 till 6;1 år) benämnde totalt sett korrekt målord i högre utsträckning jämfört med de yngre (3;9 till 4;4 år). För de yngre deltagarna var den genomsnittliga poängen 36 och för de äldre var det 45 poäng. I genomsnitt fick de äldre deltagarna nio poäng mer än de yngre.

Gemensamt för hela gruppen. Likheter som återfanns i hela åldersspannet var bland annat att endast två av nio deltagare kunde benämna orden *lätt* och *korsord*. Istället för *lätt* svarade flertalet *små* eller *liten*, och istället för *korsord* svarade flera *spel*. Vidare benämnde enbart fyra deltagare ordet *fråga* korrekt, och endast tre deltagare svarade rätt på orden *bo* och *gilla*. Ingen av deltagarna kunde benämna ordet *kräsen*.

Karaktäristiska drag för respektive åldersgrupp. De likheter som går att urskilja hos den yngre gruppen är att tre av fem svarade fel på *hel*, *kväll*, *mycket*, *ropa* och *tänka*. Ordet *hel* ersattes av två med *rund*, och två deltagare sa *stor* istället för *mycket*. Ordet *ropa* benämnde tre av fem felaktigt, varav två sa *skrika* istället och en benämnde det som *säga* (även en deltagare i det äldre åldersspannet ersatte *ropa* med *skrika*). Ordet *tänka* benämnde tre av de yngre deltagarna initialt med *gör*. Därutöver svarade samtliga deltagare i detta åldersspann inkorrekt på ordet *dag*, vilket flera benämnde som *sol*. När det kommer till deltagarna i den äldre åldersgruppen går det även där att finna gemensamma drag. Två av fyra ersatte exempelvis ordet *kräsen* med *allergisk*. För de yngre deltagarna fanns inget gemensamt svar för detta ord utan svaren varierade mellan *full*, *gillar inte det*, *inte hungrig*, *stopp* och *tåler inte*.

Frekvens. Vid jämförelse mellan de låg- och högfrekventa orden i testet kunde ingen tydlig skillnad utläsas i deltagarnas prestation. Detta gäller både för hela gruppen och för respektive åldersgrupp. Däremot fanns en skillnad inom de lågfrekventa orden där orden med allra lägst frekvens benämndes inkorrekt i högre utsträckning än orden med något högre frekvens. Ingen av deltagarna kunde benämna *kräsen*, som var det minst frekventa ordet, medan samtliga deltagare hade förvärvat ordet *moln*, som hade något högre frekvens.

Domäner. Deltagarnas svar i benämningstestet analyserades även utifrån antalet korrekta svar som deltagarna gav för de ingående orden i domänerna. En översiktlig bild av deltagarnas etablering av domänerna, med utgångspunkt i testresultatet, ges i tabell 6. Etablering av en domän innebär hur väl deltagarna behärskar de ord i domänen som testas. Observera att alla domäner inte testades.

Tabell 6. Domänernas grad av etablering hos hela gruppen utifrån testresultatet.

Antal korrekta svar för de ingående orden i respektive domän*	Domäner
9 av 9 rätt	Kroppsdel, känsla, människa, sömn, temperatur, väder
7 eller 8 av 9 rätt	Natur, storlek

5 eller 6 av 9 rätt	Form & gestalt, förbrukning, kognition, läge & rörelse, mängd, perception, ålder
4 eller färre av 9 rätt	Hushåll, kommunikation, omdöme, socialt beteende & interaktion, syssla, tid, vikt

* Baserat på det ord med minst antal korrekta svar inom en domän

Indelningen av domänerna i kolumnerna ovan har gjorts utifrån det ord, inom en domän, som erhöll minst antal korrekta svar. Exempelvis, om fem ord testades inom en domän, och samtliga deltagare svarade rätt på fyra av orden, räcker det att ett ord benämndes felaktigt av fem deltagare eller fler för att domänen ska placeras i raden *4 eller färre av 9 rätt* i tabellen ifråga.

Vad gäller de högfrekventa orden går det att avläsa i testresultatet att inom följande domäner gav samtliga deltagare korrekta svar på alla ord som testades: *kroppsdel, känsla, människa, socialt beteende och interaktion, sömn* och *temperatur*. Dock testades endast ett ord inom respektive ovannämnd domän. Flest inkorrekta svar återfanns inom domänerna *hushåll, kommunikation, omdöme, tid* och *vikt*.

Om även de lågfrekventa orden tas med är domänen *hushåll* fortsatt inte helt etablerad då flera av deltagarna svarade fel på *bädda*. Vidare bedöms domänen *socialt beteende och interaktion* likaså inte vara helt utvecklad, då ingen av deltagarna kunde benämna ordet *kräsen*. Domänerna *förbrukning, syssla* samt *väder* innehöll endast lågfrekventa ord, varav de två förstnämnda inte verkar vara fullständigt etablerade då endast ett fåtal deltagare kunde benämna de berörda orden. Domänen *väder* är däremot utvecklad i viss mån då samtliga deltagare kunde benämna ordet *moln*.

Ordklasser. Vid jämförelse mellan ordklasserna i testresultatet går det att utläsa att substantiv erhöll flest korrekta svar, beräknat i procent. Detta gäller för såväl hela gruppen som för respektive åldersgrupp.

Tabell 7. Andel korrekta svar inom ordklasserna för hela gruppen samt för respektive åldersgrupp, anges i procent.

Ordklass	Andel korrekta svar för hela gruppen	Andel korrekta svar för äldre gruppen	Andel korrekta svar för yngre gruppen
Substantiv	85,4	96,9	76,3
Verb	75,4	85,5	67,4
Adjektiv	80,2	87,5	74,3
Adverb	66,7	100	40

Vid analys av testsvaren för hela gruppen framgår det att nästintill samtliga deltagare svarade korrekt på substantiven med undantag för orden *dag* och *korsord*. Många av adjektiven besvarades likaså korrekt, med undantag för *lätt* och *kräsen*, vilka majoriteten benämnde felaktigt. I kontrast till substantiven var verb den ordklass som deltagarna hade flest uteblivna eller felaktiga svar inom. Notera att endast ett adverb (*mycket*) testades, vilket erhöll totalt sex korrekta svar inom hela gruppen. När testsvaren för respektive åldersgrupp

tas i beaktande går det även att utläsa att de äldre deltagarna svarade korrekt i högre utsträckning jämfört med de yngre. Detta gäller för samtliga ordklasser.

Sammanfattning av resultatet

För att summera ges i följande stycke en sammanställning av resultatet. Avseende frekvensordlistan bedömdes 93 innehållsord vara högfrekventa i korpusen. Dessa ord hade en relativt jämn fördelning mellan ordklasserna, varav de mest frekventa utgjordes av verb. Vidare hade 65 av dessa ord en god spridning över böckerna. De 93 högfrekventa innehållsorden delades därefter in i 27 semantiska domäner. De tre största domänerna var *läge och rörelse* (tolv ord), *natur* (nio ord) och *omdöme* (åtta ord).

Efter att de mest frekventa innehållsorden och dess tillhörande domäner fastställts valdes ett antal av de högfrekventa orden ut till ett benämningstest. Dessa ord testades tillsammans med ett litet antal lågfrekventa innehållsord. Utifrån testresultatet framkommer det att substantiv erhöll flest korrekta svar, vid jämförelse mellan ordklasserna. I kontrast till detta var verb den ordklass som deltagarna gav flest felaktiga eller uteblivna svar inom. Vid jämförelse mellan ordfrekvens i korpusen och antal korrekta svar från deltagarna kunde inget tydligt samband utläsas, det vill säga antal korrekta svar ökade inte med ökad frekvens.

Diskussion

Frekvensordlista

Ordklasser. Den sammanställda listan över de högfrekventa innehållsorden i korpusen utgörs av 93 ord. De 15 mest frekventa innehållsorden består av tolv verb, två adjektiv och två adverb (varav ett även kan klassas som adjektiv). Verben karaktäriseras av att de har en generell betydelse och av denna anledning kan de användas i många kontexter. Detta kan vara ett skäl till att de är högfrekventa och att de har en jämn fördelning över böckerna. Likaså har även många av verben som är mindre frekventa, i förhållande till ovannämnda verb, också en icke-specifik betydelse. Avseende adjektiv och adverb finns *lite/liten*, *mycket*, *stor* och *lilla* bland de mest högfrekventa, vilka alla faller inom domänerna *storlek* och/eller *mängd*. Noterbart är att dessa fyra ord står, utifrån det strukturalistiska synsättet, i paradigmatisks relation till varandra. Ordet *lite* kan sägas vara motsats till *mycket*, och *liten* kan vara motsats till *stor*. Orden befinner sig således inom samma betydelsekategori (Paradis, 2012).

Vad gäller substantiven, finns de generellt längre ner på frekvensordlistan. Detta kan förklaras av att det finns fler unika ordtyper av substantiv jämfört med verb (Goodman et al., 2008). Det större antalet unika ordtyper medför att ett givet substantiv kan användas i ett mer begränsat antal kontexter, jämfört med verb, samt att det finns ett större utbud av ord att välja mellan inom denna ordklass. Det gör det logiskt att frekvensordlistan utgörs av fler verb än substantiv samt att de representerade substantiven återfinns i lägre frekvens jämfört med de representerade verben. Samma förklaring kan ges till varför orden med sämre spridning innehåller fler substantiv (till exempel *blad*, *spöke*) än vad som återfinns bland de väldistribuerade orden.

Domäner. De högfrekventa innehållsorden kan fördelas i åtminstone 27 semantiska domäner. Den domän som innehåller flest ord är *läge och rörelse*, vilken enbart innehåller

verb. En anledning till att domänen ifråga är den största kan vara att böcker är mer dekontextualiserade jämfört med konversationer (Montag et al., 2015), vilket ställer krav på utförligare beskrivningar om vad som sker. *Läge och rörelse* är också relativt bred i sin betydelse och därför tillåter det fler ord att inkluderas än exempelvis domänen *svårighetsgrad*. Domänerna som representeras av enbart ett ord har många gånger snävare betydelser och därför begränsas antalet ord som kan inkluderas. En anledning till att ord som har en snäv betydelse ändå kan vara högfrekventa i korpusen kan vara för att de återfinns många gånger i ett litet antal böcker. En förklaring till varför *djur* och *natur* är bland de mest representerade domänerna kan således bero på att orden inom dessa domäner används flitigt i ett fåtal böcker. Det får konsekvensen att orden blir frekventa i korpusen. Däremot får de en dålig spridning över böckerna eftersom ordens snäva betydelser begränsar vilka kontexter de kan användas i.

Värt att notera är att det finns ett kulturellt inflytande vad gäller de identifierade högfrekventa innehållsorden och de domäner de tillhör. De flesta böcker i urvalet är skrivna av svenska författare och flera av böckerna reflekterar den svenska kulturen, såsom *djur* och *natur*. Flera av substantiven betecknar, till exempel, företeelser i naturen, såsom *vatten*, *träd* och *skog*. Vidare är ordet *grön* högfrekvent i korpusen, vilket också kan avspegla ämnet *natur*. I enlighet med detta menar Håkansson (2014) att ordförrådet kan variera mycket mellan individer då lexikonet bygger på sådant som barnet exponeras för. Hon påpekar däremot att det går att urskilja likheter i ordförråden hos barn som växer upp i samma kultur. Det är därmed sannolikt att om benämningstestet hade utförts på barn som är uppväxta i andra länder hade de troligtvis inte kunnat benämna ord som är relaterade till den svenska kulturen i samma utsträckning, jämfört med barn uppväxta i Sverige.

Vid jämförelse med en studie av Fenson et al. (1994), som undersökte vilka domäner de 50 första orden tillhör, kan vissa likheter urskiljas mellan de identifierade domänerna i den studien och de högfrekventa orden i uppsatsen. Mer specifikt är domänerna *människor*, *kroppsdelar*, *djur*, *hushåll*, *mat* och *fordon* gemensamma. Inom domänen *människa* ska det nämnas att orden *mamma* och *pappa* identifieras i Fensons et al. studie, vilka även finns med bland de högfrekventa orden. De valdes dock bort i uppsatsen på grund av att egennamn exkluderades. Inom domänen *mat* förekommer det överordnade begreppet *mat* i uppsatsen medan i Fensons et al. studie är orden inom den här domänen på basnivå eller underordnad nivå (*banan*, *juice*, *kaka*, *kex*, *äpple*, *ost*). Liknande mönster går att utläsa i domänen *djur*. Den här skillnaden kan förklaras av att barn lär sig ord på överordnad nivå senare i utvecklingen (omkring tre års ålder) jämfört med ord på basnivå och underordnad nivå (Nettelbladt, 2007).

Pilottestning

När det gäller testresultatet, kan de äldre deltagarna i högre utsträckning korrekt benämna målorden jämfört med de yngre. Deltagarna i den yngre åldersgruppen har i genomsnitt 36 poäng medan de äldre deltagarna fick i genomsnitt 45 poäng. De äldre har således i genomsnitt nio poäng mer än de yngre deltagarna. Det här beror rimligen på att ju äldre man är desto fler ord har man exponerats för, både i tal och text.

Karaktäristiska drag för respektive åldersgrupp. I testresultatet kan likheter urskiljas för såväl gruppen i sin helhet som inom de två åldersgrupperna. Likheter som går att finna bland de fem yngre deltagarna är att majoriteten av dessa har svårt att benämna ordet

tänka, vilket ofta ersätts med *gör*. Vidare ersätts ordet *ropa* med *skrika* eller *säga* av flera deltagare. Samtliga tre ord (*skrika*, *ropa*, *säga*) befinner sig dock inom samma domän (*kommunikation*). Såväl *tänka* som *ropa* ersätts följaktligen ofta av verb som har en mer generell betydelse, vilket troligtvis kan bero på att dessa domäner (*kognition* respektive *kommunikation*) inte är fullständigt etablerade.

Andra gemensamma drag som finns hos deltagarna i det yngre åldersspannet är att de tenderar att utgå mer från bildstödet, snarare än meningskontexten, jämfört med de äldre deltagarna. *Dag* och *hel* ersätter flera med *sol* respektive *rund*. Ett annat exempel är när *liten* benämns som *röd*. Orden som de ersätter målorden med passar till viss del med bilden men inte alls med handlingen. Att de yngre barnen utgår ifrån bildstödet i högre utsträckning än de äldre kan troligtvis dels förklaras av att det är kognitivt krävande att förstå meningssifyllnad. Ett annat skäl kan vara att bilderna är ett starkt visuellt stimuli, vilket fångar de yngre barnens uppmärksamhet mer än den auditiva signalen. Det ska dock nämnas att det i båda grupperna även finns flera exempel på när deltagarna utgår mer från meningskontexten än bildstödet. Exempelvis svarar en deltagare *problem* istället för *korsord* som respons till meningen "...och löser ett...". En anledning till det kan emellertid vara att *löser ett problem* är en lexikaliserad fras. Slutligen kan endast en likhet, som inte går att finna hos de yngre, identifieras bland de äldre deltagarna. Två av dessa fyra deltagare benämner *kräsen* som *allergisk*. Ordet *allergisk* är förmodligen ett ord som inte finns i de yngre barnens lexikon.

Gemensamt för hela gruppen. De likheter som går att se inom hela gruppen är bland annat att flera överanvänder ordet *stor*, som flertalet gånger ersätter orden *gammal*, *hel*, *lång*, *lätt*, *mycket* och *stå*. Detta kan delvis förklaras av att *stor* är det ord bland de dimensionella termerna som barn tillägnar sig först (Clark, 1972). I korpusen går det också att utläsa att barn exponeras för *stor* i hög utsträckning. Vidare framkommer, för nästintill alla deltagare, svårigheter att benämna orden *bo*, *gilla* och *fråga*. Det här kan sannolikt förklaras av att dessa ord är svåra att elicitera, vilket kan bero på att de är svårillustrerade och/eller har en bristfällig placering i meningskontexten. Således kan man inte med säkerhet avgöra i vilken uträkning dessa ord är förvärvade hos deltagarna eftersom eliciteringsmetoden eventuellt kan innehålla felkällor. Noterbart är även att felsvaren som deltagarna ger ofta är möjliga och relevanta i sammanhanget och det är sällan deltagarna inte ger något svar alls.

En annan likhet som går att se hos flera är att de svarat inkorrekt på ordet *lätt*, genom att de flesta har ersatt ordet med *små* eller *liten*. Många av deltagarna tar även hjälp av meningskontexten och motsatsordet *tung* när de ska benämna *lätt*, då de svarar *inte tunga* som respons till "*de ser inte ut att vara tunga utan väldigt...*". Det här indikerar således att barnen inte har etablerat domänen *vikt* fullständigt. Det ska dock upprepas att i testet avsågs *vikt* men ordet *lätt* kan även syfta på *svårighetsgrad*. Resultatet kan ha påverkats av att *lätt* med betydelsen *vikt* inte åsyftas lika frekvent som betydelsen *svårighetsgrad* i korpusen, vilket upptäcktes först efter att testet genomförts. Ett annat skäl till att deltagarna ofta svarade fel på ordet *lätt* kan ges av Gärdenfors (2014) teori om utvecklingen av semantiska domäner. Enligt honom utvecklas domäner genom att de separeras från varandra. För ordet ifråga förefaller flera av barnen inte ha separerat domänen *vikt* från *storlek* eftersom flera svarade *små* eller *liten* istället för *lätt*. Samma förklaring kan ges till varför två av deltagarna använder ordet *stor* för *mycket*.

Vidare går det hos samtliga deltagare att utläsa att ordet *kräsen* inte är förvärvat då ingen kan producera det vid testning. Majoriteten av deltagarna benämner även *korsord* felaktigt. Det här resultatet kan förklaras av att dessa två ord är lågfrekventa, vilka lärs in senare jämfört med högfrekventa ord (Hansen, 2017; Leonard, 1998). Ett annat skäl till att deltagarna inte har tillägnat sig *kräsen* och *korsord* kan möjligtvis även bero på att dessa ord är ålderdomliga, som i sin tur kan vara en anledning till att de är lågfrekventa i korpusen.

Hansen (2017) genomförde en studie som visade att frekvensen har stor betydelse, i barnriktat språk, för vid vilken tidpunkt barn tillägnar sig ett ord. Däremot stämmer inte detta antagande med de resterande lågfrekventa målorden då flertalet av deltagarna kan benämna dessa. En anledning till det kan emellertid vara att dessa ord är mer frekventa i tal än i böcker, alternativt att korpusen inte är representativt. I kontrast till detta anser Bloom (2000) att endast en exponering kan vara tillräcklig för att ett barn ska tillägna sig ett nytt ord. Frekvensen behöver därför inte vara avgörande för hur lätt ett ord lärs in. Ytterligare menar Gärdenfors (2014) att det är lättare att förvärva ett nytt ord i en domän som redan är etablerad. Det kan förklara varför endast en exponering kan räcka för att barn med typisk språkutveckling ska lära sig ett nytt ord.

Frekvensens betydelse. Det går inte att utläsa någon god överensstämmelse hos de undersökta orden mellan ordets totala förekomst och hur många barn som kunde benämna det. Det vill säga, antalet korrekta svar ökar inte med ökad frekvens. Något sådant mönster går varken att se inom hela gruppen eller i någon av åldersgrupperna. Det här säger emellertid inte emot frekvensens betydelse för ordinlärning eftersom majoriteten av barnen kan benämna nästintill alla högfrekventa ord. Ytterligare är många av de högfrekventa verben svåra att testa, vilket troligtvis påverkar resultatet.

Etablerade domäner. Med utgångspunkt i testresultatet framkommer det att följande domäner förefaller vara etablerade: *kroppsdel, känsla, människa, sömn, temperatur* och *väder*. Domänerna *natur* och *storlek* förefaller också vara någorlunda utvecklade då de erhöll endast ett fåtal felsvar. Mest inkorrekta svar återfinns inom domänerna *hushåll, kommunikation, omdöme, socialt beteende och interaktion, syssla, tid* samt *vikt*. Några av felsvaren inom domänerna *hushåll, kommunikation, läge och rörelse* samt *omdöme* beror emellertid sannolikt på felkällor i metoden. I till exempel domänen *läge och rörelse* förekommer flera felsvar för ordet *åka*, vilket kan bero på ett otydligt bildstöd, medan resterande ord i domänen benämns korrekt av nästintill alla deltagare. Domänen förefaller därför vara etablerad i relativt hög utsträckning. Att domänerna *socialt beteende och interaktion* samt *syssla* erhöll många felsvar beror troligtvis på att ett eller flera av deras ingående ord är lågfrekventa. *Kräsen* (som ingår i domänen *socialt beteende och interaktion*) är även ett abstrakt ord, vilket kan bidra till att det blir svårare att förvärva (Hansen, 2017).

Ytterligare innehåller domänen *perception* flera felsvar och förefaller därmed inte vara fullständigt etablerad. Men med tanke på att felsvaren är relativt avancerade och specifika synonymer (till exempel *kika* och *speja*) till målorden (*se* och *titta*) tycks emellertid domänen vara någorlunda utvecklad. Avslutningsvis bör det återigen framhållas att endast ett eller ett fåtal ord inom respektive domän testats och därmed går det inte att med säkerhet avgöra huruvida en domän är etablerad.

Verbens komplexitet. I frekvensordlistan framkommer det att verb är den ordklass som är överlägset mest högfrekvent. Testresultatet visar däremot samtidigt att denna ordklass

innehåller flest felsvar, såväl inom hela gruppen som inom respektive åldersgrupp. Skälen till detta kan vara att de är svåra att testa samt att substantiv många gånger exemplifieras med en tydlig referent, till skillnad från verb. Vuxna i barnets omgivning benämner ofta föremål medan handlingar vanligtvis inte utförs samtidigt som de betecknas (Gleitman, 1990; Tomasello, 1992). Hansen (2017) resonerar även att ord som är lättare att skapa en mental representation av förvärvas tidigare. Det här kan följaktligen vara ännu en anledning till att deltagarna ger fler korrekta svar på substantiv i jämförelse med verb.

Metoddiskussion

Eliciteringsstrategier. Valet att utforma benämningstestet som en saga motiveras av att ord lärs in i en meningssammanhang, vilket är i enlighet med synsättet inom kognitiv semantik (Paradis & Willners, 2011). Barn exponeras sällan för ett ord isolerat utan hör det tillsammans med andra ord (Bloom, 2000). Den valda metoden är därför mer ekologiskt valid jämfört med konfrontationsbenämning, som hade varit en alternativ testmetod. Den omgivande kontexten ger ledtrådar till det avsedda ordets betydelse och därmed är den lexikala betydelsen till stor del avhängig vilket sammanhang ordet används i (Paradis & Willners, 2011). Exempelvis är ordet *år* för sig självt svårt att elicitera men blir lättare att locka fram genom en kontext. Eftersom ordet ofta förekommer tillsammans med *fylla* utnyttjades detta språkliga sammanhang vid testningen. Däremot kan den här kontexten samtidigt tolkas som en för lätt ledtråd då *fylla år* är en välkänd lexikaliserad fras. Möjligen behärskar således barn *fylla år* som lexikaliserad fras, men inte orden var för sig. Enligt testresultatet kan samtliga deltagare ordet *år*, vilket sålunda kan ha påverkats av den meningssammanhang som ordet står i. Ett annat exempel på en lexikaliserad fras i benämningstestet är *mitt i natten*.

Vidare blir verb lättare att elicitera genom en saga då dessa ofta är svårillustrerade. Ordet *hjälpa* hade till exempel blivit svårt att få deltagarna att säga genom konfrontationsbenämning, men eftersom det förekom i en kontext förstod de flesta barn vilket ord som åsyftades. Det ska dock nämnas att några ord som inkluderades i testet är svåra att elicitera trots att de förekom i en mening. Orsaken till att ett målord kunde vara svårt att locka fram kan antingen bero på att valet av bildstöd var bristfälligt, till exempel *åka* (som föreställer en person som kör bil), eller för att strategin för meningsifyllnad inte var väl avvägd. Bristfällig meningsifyllnad innebär i det här sammanhanget att orden som föregår målordet inte är tillräckligt tydliga för att kunna elicitera det, till exempel: *deras sång väcker mamma som fortfarande... (ligger) i sängen*. I de fall då deltagarna benämnde fel målord vid meningsifyllnad kunde de dock ofta producera korrekt ord genom att testledaren tillämpade frågestrategi, till exempel: *vilket sätt vilar hon på?* Sammantaget fungerade testförfarandet relativt väl, både vad gäller sagan och bildstödet.

Förslag på kompletterande test. Trots att många ord är lättare att elicitera genom en saga exkluderades flera ord som ansågs vara alltför ospecifika från testet, exempelvis orden *göra* och *vilja*. Ett sätt att testa även dessa ord kan vara att utföra ett förståelsetest (till exempel utpekning av bilder med ett antal semantiska distraktorer), alternativt be deltagarna att beskriva orden. Förutom att testa de svårillustrerade orden hade det även varit intressant att undersöka om deltagarna förstod målorden som de inte kunde benämna. Som tidigare nämnts förstår barn fler ord än de kan producera (Clark, 1993), vilket beror på att produktion är mer krävande. Vid ordproduktion behöver barnet hämta lagrade representationer och relevant

artikulatorisk information, liksom utföra lämpliga talmotoriska rörelser (Clark, 2009). Ett kompletterande förståelsetest kan ge information om vilka ord som är inom barnets proximala utvecklingszon. Goldin-Meadow et al. (1976) undersökte tvååringars förmåga till ordproduktion och ordförståelse. De fann att barnen förstod fler ord än de kunde producera och denna skillnad var större för verb jämfört med substantiv. Genom att göra ett förståelsetest hade man kunnat få insikt i om denna skillnad även gäller för äldre barn.

Tidsåtgång. Utförandet av testet tog 15 till 20 minuter, beroende på hur svårt deltagaren hade för att benämna målorden. En av deltagarna blev ofokuserad i slutet vilket kan ha påverkat resultatet eftersom denna person benämnde de sista sju målorden felaktigt. Övriga deltagare uppvisade emellertid inga tecken på rastlöshet, vilket indikerar att testet är lagom långt. Möjligtvis hade det kunnat kortas ner cirka tre minuter för att samtliga deltagare skulle kunna koncentrera sig optimalt. Alla kunde dock fullfölja testningen.

Begränsningar. Vidare går det inte med säkerhet att avgöra med den tillämpade metoden om en domän är etablerad eftersom endast enstaka ord inom respektive domän testas. På så sätt är det eftersträvansvärt att inkludera fler ord från respektive domän men urvalet möjliggjorde inte detta. Det hade även påverkat tidsåtgången. Topplistor från fler bibliotek hade också varit önskvärt med syfte att få större spridning på böckerna. För att kunna generalisera och dra välgrundade slutsatser krävs dessutom ett större urval av såväl böcker som deltagare. Ytterligare en begränsning är att frågestrategi inte utfördes på ett strukturerat sätt vilket medförde att deltagarna fick olika frågor, och därmed olika mängd stöd, vid elicitering av målord. Uppsatsförfattarna har heller inte tillfrågat vårdnadshavarna hur ofta barnen blir lästa för. Gemensam bokläsning är, som nämnts, en viktig källa för ordförrådstillväxten. Barn som läser fler böcker med sina vårdnadshavare uppvisar bättre språkförmågor, däribland ett större lexikon, än de barn som läser färre böcker (Bus, van Ijzendoorn & Pellegrini, 1995; Farrant & Zubrick, 2012; Payne, Whitehurst & Angell, 1994).

Slutligen finns en begränsning i beräkandet av ordfrekvenser. Ordet *lite* var till exempel högfrekvent i korpusen men ingen hänsyn togs till det faktum att böjningsformerna *litet* och *liten* inte tillhör samma ordklass då *lite* är ett adverb medan *litet/liten* är ett adjektiv. Även *lilla* och *små*, som analyserats separat, är egentligen också andra böjningsformer av adjektivet *litet/liten*. I benämningstestet inkluderas adjektivet men inte adverbet. Därmed går det inte genom testresultatet att uttala sig om huruvida adverbet ifråga är förvärvat.

Slutsatser

Antalet identifierade högfrekventa innehållsord är 93, varav majoriteten av de mest högfrekventa utgörs av verb. Karaktäristiskt för flertalet verb är att de har en generell betydelse. Av de 93 orden återfinns flest ord inom de semantiska domänerna *djur, läge och rörelse, natur, omdöme, tid* och *tillstånd*. Vidare identifierades 65 av de högfrekventa orden ha god spridning över böckerna, varav 43 inkluderades i ett benämningstest. Utifrån testresultatet förefaller barn i åldrarna tre till sex år ha förvärvat mer än hälften av orden i hög utsträckning och resterande ord i viss omfattning. Inget tydligt samband går att utläsa mellan ords frekvens i korpusen och antal barn som kan benämna dem.

Implikationer för logopediskt arbete och framtida forskning

I dagsläget finns begränsat med kunskap gällande vilka ord och semantiska domäner barn förväntas behärska vid olika åldrar. Resultatet kan därmed bidra till större insikt kring den lexikal-semantiska utvecklingen hos enspråkiga barn. Vidare kan det ge vägledning kring vilka målord och semantiska domäner logopedier bör ha som utgångspunkt vid bedömning av barns lexikal-semantiska förmåga samt vid intervention för barn med språkstörning. Ytterligare framhåller testresultatet vikten av att språksätta handlingar, det vill säga att personer i barnets omgivning sätter ord på vad barnet gör, för att underlätta tillägnande av verb.

Studien kan fungera som underlag för utformandet av ett benämningstest för bedömning av ordförrådet. Genom att konstruera ett test som en saga blir testsituationen mindre styrd, jämfört med konfrontationsbenämning, och det underlättar elicitering av målord. Testningen blir dessutom mer ekologiskt valid eftersom ord lärs in i en meningssammanhang och inte isolerat. Det är därmed motiverat att utveckla fler benämningstest med utgångspunkt i hur ord faktiskt lärs in. Uppsatsens tillvägagångssätt i skapandet av en korpus samt benämningstest kan således fungera som inspirationskälla för framtida forskning inom området.

TACK

Ett varmt tack riktas till alla barn som deltagit i testningen samt till förskolepedagogerna för er hjälp i rekryteringsprocessen. Vi vill även ägna ett stort tack till Ketty Andersson och Kristina Hansson för er goda handledning och för det stöd vi känt under hela arbetsgången.

Referenser

- Anthony, L. (2018). AntConc (Version 3.5.8) [Datorprogram]. Hämtad från <http://www.laurenceanthony.net/software/antconc/>.
- Barrett, M. (1995). Early lexical development. I P, Fletcher., & B, MacWhinney. (Red.), *The handbook of child language*. 325–343. Oxford: Basil Blackwell.
- Blackwell, A. (2005). Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology. *Journal of Child Language*, 32, 535–562. <http://dx.doi.org.ludwig.lub.lu.se/10.1017/S0305000905006938>.
- Bloom, P. (2000). *How children learn the meanings of words*. London: The MIT Press Cambridge.
- Bus, A. G., van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65(1), 1–21. <https://www-jstor-org.ludwig.lub.lu.se/stable/1170476>.
- Clark, E. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning and Verbal Behavior*, 11(26), 750–758. [https://doi.org/10.1016/S0022-5371\(72\)80009-4](https://doi.org/10.1016/S0022-5371(72)80009-4).
- Clark, E. (1993). *The lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, E., & Wong, A. (2002). Pragmatic directions about language use: Words and word meanings. *Language in Society* 31(2), 181–212. <https://www-jstor-org.ludwig.lub.lu.se/stable/4169168>.
- Clark, E. V. (2009). *First language acquisition* (2. uppl.). Cambridge: Cambridge University Press.
- Cruse, A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cruse, D.A. (2002). *The construal of sense boundaries*. *Revue de Sémantique et Pragmatique*, 12, 101–119.
- Darts kommunikation- och dataresurscenter. (u.å). Bildstöd – Kommunikationsstöd i vårdssituationer. Hämtad 2019-02-26 från <http://bildstod.se>.
- De Geer, S & Gunnarsdottir, C. (2013). *Om en person inte är gammal...så är han inte så himla ruskig i ansiktet!* (Magisteruppsats, Lunds universitet, Lund). Hämtad från

<http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=3809730&fileOId=3809734>.

- Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, 32(3), 343–364. <https://doi.org/10.1177/0142723711422626>.
- Fellbaum, C. (1998). *WordNet: An Electronic lexical database*. Cambridge: The MIT Press.
- Fenson, L., Dale, P. S., Reznick, J., Steven, Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development* 59(5). DOI: 10.2307/1166093.
- Ganidou, E. (2007). *Det var en gång en bilderbok - en semiotisk undersökning av två bilderböcker*. (Examensarbete, Malmö Högskola, Malmö). Hämtad från <http://muep.mau.se/bitstream/handle/2043/6573/Examensarbete2007.pdf?sequence=1&isAllowed=y>.
- Gleitman, L.R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55. http://doi.org/10.1207/s15327817la0101_2.
- Goldin-Meadow, S., Seligman, M. E. P. & Gelman, R. (1976). Language in the two-year-old. *Cognition*, 4, 189–202. [https://doi.org/10.1016/0010-0277\(76\)90004-4](https://doi.org/10.1016/0010-0277(76)90004-4).
- Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, 35, 515–531. DOI: 10.1017/S0305000907008641.
- Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge: MIT Press.
- Gärdenfors, P. (2018, januari). From Sensations to Concepts: a Proposal for Two Learning Processes. *Review of Philosophy and Psychology*, 1–24. <https://doi.org/10.1007/s13164-017-0379-7>.
- Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*. 37(2), 205-225. DOI: 10.1177/0142723716679956.
- Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, 73(2), 418–433. <https://www-jstor-org.ludwig.lub.lu.se/stable/3696366>.
- Håkansson, G. (2014). *Språkinläring hos barn* (2. uppl.). Lund: Studentlitteratur.

- Johansson Kokkinakis, S. (2007). I I. Lindberg & S. Johansson Kokkinakis (Red.), *OrdiL – en korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. (ROSA 8) Göteborg: Göteborg: Institutet för svenska som andraspråk, Göteborgs Universitet.
- Järborg, J. (2007). Om ord och ordkunskap. I I. Lindberg & S. Johansson-Kokkinakis (Red.), *OrdiL – en korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. (ROSA 8). Göteborg: Institutet för svenska som andraspråk, Göteborgs Universitet.
- Lehrer, A. (1985). The Influence of Semantic Fields on Semantic Change. I J. Fisiak. (Red.), *Historical semantics – Historical Word-Formation*, 283–296. Berlin: Mouton.
- Leonard, L. (1998). *Children with specific language impairment*. Cambridge, Mass: MIT Press.
- Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. I C. Gallaway & B.J. Richards (Red.), *Input and interaction in language acquisition*, 56-73. Cambridge: Cambridge University Press.
- Luniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anelkovic, D., Unal-Logacev, O. (2015). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, 48(3), 1154–1177. DOI: 10.3758/s13428-015-0636-6.
- Miller, G. A. (1996). *The science of words*. New York: Freeman.
- Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The Words Children Hear: Picture Books and the Statistics for Language Learning. *Psychological Science*, 26(9), 1–8. DOI: 10.1177/0956797615594361.
- Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications of acquisition and instruction. I M.G McKeown & M.E Curtis (Red.), *The nature of vocabulary acquisition*. Hillsdale, NJ: Erlbaum.
- Nationalencyklopedin. (u.å). Ord. Hämtad 2019-01-29 från <http://www.ne.se/uppslagsverk/encyklopedi/lång/ord>.
- Nationalencyklopedin. (u.å). Semantik. Hämtad 2019-02-12 från <https://www-ne-se.ludwig.lub.lu.se/uppslagsverk/encyklopedi/lång/semantik>.
- Nettelbladt, U. (2007). Lexikal utveckling. I U. Nettelbladt & E. K. Salameh (Red.), *Språkutveckling och språkstörning hos barn*. Lund: Studentlitteratur.

- Paradis, C. (2003). Is the notion of linguistic competence relevant in Cognitive Linguistics? *Annual Review of Cognitive Linguistics*, 1(1), 247–271.
[https://portal.research.lu.se/portal/sv/publications/is-the-notion-of-linguistic-competence-relevant-in-cognitive-linguistics\(403613f9-3996-46b0-acbb-3d90b08b2a15\).html](https://portal.research.lu.se/portal/sv/publications/is-the-notion-of-linguistic-competence-relevant-in-cognitive-linguistics(403613f9-3996-46b0-acbb-3d90b08b2a15).html).
- Paradis, C., & Willners, C. (2011). Antonymy: from conventionalization to meaning-making. *Review of Cognitive Linguistics*, 9(2), 367–391.
<https://lup.lub.lu.se/search/publication/367f87f8-5ef4-4663-b84d-1ccf8dafa432>.
- Paradis, C. (2012). Lexical Semantics. I C. A. Chapelle (Red.), *The encyclopedia of applied linguistics*. DOI: 10.1002/9781405198431.wbeal0695.
- Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly*, 9(3-4), 427–440.
[http://dx.doi.org/10.1016/0885-2006\(94\)90018-3](http://dx.doi.org/10.1016/0885-2006(94)90018-3).
- Piaget, J. (1972). *The psychology of the child*. New York: Basic Books.
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83(5), 1762–1774. DOI: 10.1111/j.1467-8624.2012.01805.x.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Weisleder, A., & Fernald, A. (2014). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, 24(11), 2143–2152. DOI: 10.1177/0956797613488145.
- Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, 37(2), 265–279. DOI: 10.1037//0012-1649.37.2.265.

Bilaga 1 - Böcker

- Adbåge, E. (2018). *Gropen*. Stockholm: Rabén & Sjögren
- Adolfsson, E. (2005). *Den blå barnkammARBOKEN*. Stockholm: Bonnier Carlsen
- Anderson, L & Svedberg, U. (1983). *Maja tittar på naturen*. Stockholm: Rabén & Sjögren
- Bergström, G. (1973). *Aja baja, Alfons Åberg!* (4 uppl.). Stockholm: Rabén & Sjögren
- Bergström, G. (1977). *Listigt, Alfons Åberg!* (4 uppl.). Stockholm: Rabén & Sjögren
- Bergström, G. (1982). *Var är bus-Alfons?* Stockholm: Rabén & Sjögren
- Björkvik, A & Midbeck, C. (2018). *En känsla av mig*. Jönköping: Ehrlin Publishing
- Dahlin, E & Lundeberg, A. (2006). *Den gula barnkammARBOKEN*. Stockholm: Bonnier Carlsen
- Ericsson, L. (2018). *Pruttresan: Sagan genom landet inuti*. Stockholm: Triumf Förlag
- Forssén, Ehrlin, C-J. (2011). *Kaninen som så gärna ville somna: En annorlunda godnattsaga*. Jönköping: Ehrlin Publishing
- Forssén, Ehrlin, C-J & Vintrafors, K. (2018). *Modiga Morris: En vecka upp och ner*. Jönköping: Ehrlin Publishing
- Glori, D. (2017). *Vad som än sker*. Västra Frölunda: Tukan Förlag
- Kruusval, C. (2017). *Ellens alla hundar*. Stockholm: Rabén & Sjögren
- Landström, L & Landström, O. (1999). *Bu och Bä i skogen*. Stockholm: Rabén & Sjögren
- Lindgren, A. (2017). *Räven och tomten*. Stockholm: Rabén & Sjögren
- Läckberg, C & Larsson, T. (2018). *Super-Charlie och spökfällan*. Stockholm: Bonnier Carlsen
- Nordqvist, S. (2001). *När Findus var liten och försvann*. Stockholm: Opal
- Nordqvist, S. (2007). *Pannkakstårtan*. Stockholm: Opal
- Nordqvist, S. (2012). *Findus flyttar ut*. Stockholm: Opal
- Norlin, A. (2017). *Här kommer färjan*. Stockholm: Rabén & Sjögren
- Sheppard, S. (2016). *Djuren i skogen. Vinterboken*. Stockholm: Rabén & Sjögren
- Sjögren, S & Forsmark, A. (2018). *Lilla nej-boken*. Stockholm: Libris
- Sandberg, I & Sandberg, L. (2015). *Lilla spöket Laban*. Stockholm: Rabén & Sjögren
- Wieslander, J. (2016). *Mera fart Mamma Mu*. Stockholm: Rabén & Sjögren
- Wieslander, J. (2018). *Mamma Mu låtsas*. Stockholm: Rabén & Sjögren
- Wieslander, J. (2019). *Mamma Mu simmar*. Stockholm: Rabén & Sjögren

Bilaga 2 - Högfrekventa innehållsords frekvens och distribution över böckerna

De högfrekventa innehållsorden presenteras i frekvensordning, i fallande led. Ordens fördelning över böckerna redovisas genom antal böcker som ordet förekom i någon gång, samt i hur många böcker ordet återfanns minst 2,5 procent av dess totala förekomst. I den sista kolumnen presenteras även det lägsta och högsta värdet för respektive ords absoluta frekvens i förhållande till gränsen 2,5 procent. Den lägsta frekvensen motsvarar således minst 2,5 procent. Exempelvis återfinns ordet *gå* någon gång i 24 böcker. I elva av dessa böcker förekom ordet flera gånger (minst 2,5 procent av dess totala förekomst). Den lägsta förekomsten, inom detta gränsvärde, var sju förekomster och den högsta var 44.

Tabell för respektive ords distribution i korpusen.

Ord	Ordets totala förekomst i korpusen	Antal böcker som ordet förekommer i	Antal böcker som ordet förekommer i över 2,5 procent	Lägsta och högsta förekomst (absolut frekvens) i förhållande till 2,5 procent
Vara*	1557	26	11	46-230
Ha*	623	26	12	22-112
Säga*	570	26	12	15-74
Kunna*	488	25	12	14-89
Ska*	340	25	11	5-75
Få	332	24	14	9-48
Bli*	304	24	12	9-51
Komma*	304	25	10	8-53
Gå	269	24	11	7-44
Lite/liten	265	23	9	10-50
Se	228	25	14	6-50
Göra*	214	23	13	6-40
Ta*	176	23	16	5-38
Mycket	173	21	10	5-38
Stor	171	20	9	5-36
Lilla	165	13	6	6-57
Finnas*	162	20	11	4-54

Vilja*	152	20	11	4-28
Dag	140	22	11	5-21
Sitta	137	22	10	4-28
Tänka	136	20	11	4-30
Lång	135	19	9	4-27
Djur	133	8	5	7-55
Bra	130	21	10	4-29
Vinter	130	5	3	6-81
Hel	128	18	12	4-19
Äta	123	18	7	7-34
Många	122	18	8	3-55
Vatten	119	11	8	5-54
Små	111	14	5	3-44
Fågel	102	6	4	3-75
Träd	101	14	9	2-50
Katt	97	12	7	3-26
Sova	97	19	9	3-22
Tycka*	91	19	8	3-16
Mat	84	13	6	4-28
Titta	84	22	10	3-14
Måste*	83	19	8	4-15
Börja*	81	19	15	2-12
Flyga	81	15	11	2-15
Åka	81	12	8	2-54
Ligga	80	20	13	2-13
Bil	77	7	3	4-63
Somna	77	12	6	2-58

Hus	75	13	8	3-24
År	75	10	9	2-24
Stå	74	18	13	2-16
Rund	73	15	14	2-19
Hjälpa	72	14	9	2-16
Natt	72	15	11	1-12
Hoppa	71	15	9	2-30
Ny	71	13	11	2-15
Gammal	69	12	11	2-13
Ropa	69	18	10	2-12
Skog	69	11	8	2-35
Barn	66	17	11	2-21
Snö	66	7	5	2-22
Gubbe	65	6	6	2-19
Spöke	64	5	4	3-22
Trött	64	5	2	7-55
Veta*	64	20	15	1-9
Tro*	58	13	11	2-16
Hem	56	17	13	2-10
Blad	54	3	1	52-52
Fråga	53	15	8	2-13
Leka	53	15	7	2-12
Hund	51	9	5	2-26
Mark	51	11	7	2-27
Spindel	51	8	4	3-22
Öga	51	12	8	3-9
Bo	50	13	8	2-10

Röd	50	7	5	3-21
Cykel	49	6	5	2-25
Säng	49	9	7	2-22
Glad	48	15	9	2-11
Köra	47	7	5	4-24
Riktig	47	16	9	2-9
Blomma	46	5	3	2-37
Mage	45	9	8	2-15
Rädd	44	9	7	2-19
Snäll	44	12	6	2-23
Tåg	44	4	2	2-39
Växt	42	3	2	6-35
Höna	41	7	5	3-13
Precis	41	14	7	2-10
Huvud	40	10	10	1-13
Kväll	39	14	14	1-9
Lätt	39	13	13	1-10
Varm	38	14	14	1-14
Vår	38	3	3	2-18
Gilla	37	10	10	1-14
Gräs	35	11	11	1-18
Grön	35	10	8	1-13

(*) svårillustrerade ord. Ordet "få" avser verbet och "vår" avser substantivet

Bilaga 3 - Benämningstest

Sagan om Sally och Josef

Nu ska du få höra en saga som handlar om en familj som för en kort tid sedan flyttade från stan ut till landet. Till skillnad från inne i stan är här väldigt tyst och lugnt. Det enda som hörs är fåglarna som kvittrar och suset från vinden. Fast ibland galer grannens tupp så högt att det stör den lugna miljön. Trots tuppen, trivs familjen ganska bra på landet. Här kan man bada i sjön, plocka blommor på ängarna och gosa med grannarnas kattungar.

Familjen har flyttat till ett fint rött och gult... **hus (1)**. I huset... **bor (2)** mamma, pappa och deras två... **barn (3)**, Sally och Josef. Det är mörkt ute för det är mitt i... **natten (4)**. Sch, vi måste vara tysta så vi inte väcker dom. För just nu... **sover (5)** dom i sina sängar. När solen har gått upp och natt har blivit... **dag (6)** väcker pappa Sally och Josef och berättar att det är mammas födelsedag. Dom är lediga från förskolan för att fixa ett kalas för mamma eftersom hon fyller... **år (7)**. Pappa bär på en bricka med frukost och presenter som mamma ska... **få (8)**. Dom går tillsammans in i sovrummet för att sjunga "Ja må hon leva" för henne. Deras sång väcker mamma som fortfarande... **ligger (9)** i sängen. När mamma får syn på pappa och barnen skiner hon upp i ett leende. Hon ser väldigt... **glad (10)** ut.

Medan mamma är på jobbet passar dom på att städa i huset för att göra det fint. Josef plockar upp leksaker från golvet. Sally dammtorkar i vardagsrummet. Pappa går in i sovrummen och... **bäddar (11)** sängarna. Efter att dom städat färdigt ska dom baka en tårta. Tårtan ska dom bjuda mamma på när hon kommer... **hem (12)**. Sist dom bakade med pappa råkade tårtan bli väldigt... **liten (13)**. Idag ska dom följa receptet och göra en... **stor (14)** tårta. För att tårtan ska bli stor krävs det inte lite utan... **mycket (15)** smet. När dom ska plocka fram alla ingredienser upptäcker pappa att mjölet är slut. Pappa ber Sally att hämta mer mjöl som finns nere i... **källaren (16)**. Sally vill inte hämta det själv eftersom hon tycker att källaren är... **läskig (17)**. Pappa följer därför med henne. Efter det måttar Sally upp mjöl medan Josef knäcker ägg. När dom har vispat ner alla ingredienser håller pappa ner smeten i en kakform som är... **rund (18)**. Sedan sätter pappa in formen i ugnen. Medans tårtan gräddas går barnen ut och leker. Pappa städar undan i köket. I vasken ligger äggskal som Josef lämnat kvar. Pappa suckar, tar upp äggskalerna och... **slänger (19)** dom i soporna.

Bredvid huset finns en... **skog (20)** och på husets baksida syns en massa... **gräs (21)** och ett högt... **träd (22)**. På ena trädgrenen har mamma och pappa hängt upp en gunga. Sally har spring i benen och springer genast fram till gungan. Josef följer efter men springer inte utan han... **går (23)**. Josef har lite svårt att själv komma upp på gungan så Sally... **hjälp (24)** honom. Efter att Josef gungat färdigt, är det Sallys tur. Sally gungar varje dag. Hon... **gillar (25)** att gunga. Hon tycker särskilt om att gunga högt. För att få så hög fart som möjligt... **står (26)** Sally upp på gungan. En gång gick gungan i sönder och då fick dom köpa en... **ny (27)**. När Sally gungar hör hon plötsligt ett läte som kommer ovanifrån. När hon... **tittar (28)** upp får hon syn på en fågel som kraxar och breder ut sina vingar medan den... **flyger (29)**. På himlen ser hon också flera vita, fluffiga... **moln (30)**. Förutom att gunga tycker Sally om

att cykla. Hon har precis lärt sig att cykla utan stödhjul. Till en början trillade hon ofta och därför är det viktigt att ha en hjälm på... **huvudet (31)**.

Under tiden som Sally gungar går Josef till dammen i trädgården. Dammen är fylld med... **vatten (32)** och i den simmar några fiskar. Vid dammen hittar Josef en groda, som är alldeles... **grön (33)**. Grodans... **ögon (34)** är stora och lyser gult. Josef tror att grodan... **ser (35)** bra i mörker. Josef tycker att grodan ser fundersam ut. Han undrar vad den... **tänker (36)**. Det vore häftigt att kunna läsa grodans tankar. Josef studerar grodan och ser att den tar sig fram genom att... **hoppa (37)**. Den kan hoppa både högt och... **långt (38)**. Josef hittar även en myrstack i trädgården. Det är svårt att räkna myrorna. Det finns inte en eller två eller tre utan jätte... **många (39)**. Myrorna är små. Dom ser inte ut att vara tunga utan väldigt... **lätta (40)**.

När tårtan är klar går pappa ut på trappan och... **ropar (41)** på barnen. Sally och Josef springer in i köket och får syn på tårtan i ugnen. Pappa tar ut tårtan och säger att dom inte får röra den ännu då den är väldigt... **varm (42)**. Nu när tårtan är färdig väntar dom bara på att mamma ska komma hem så att kalaset kan börja. Under tiden leker Sally och Josef med lego medan pappa... **sitter (43)** i fätöljen och löser ett... **korsord (44)**. När det blivit eftermiddag och mamma är klar på jobbet så... **åker (45)** hon hemåt. I dörren möts hon av Sally, Josef och pappa. Barnen springer fram och kramar mamma. Sally har glömt hur många år mamma fyller så hon... **frågar (46)** henne. Mamma svarar att hon blir 38 år. Oj så mycket tänker Sally. Mamma börjar verkligen bli... **gammal (47)**. Sedan är det äntligen dags att äta tårtan. Alla tycker att den är jättegod. Det tycker till och med Josef, som annars inte gillar så många sorters mat, eftersom han är... **kräsen (48)**. Efter att dom har fikat färdigt har en stor del av tårtan blivit uppäten. Nu är den halv, innan dess var den... **hel (49)**.

När klockan slagit sju är det dags för Sally och Josef att borsta tänderna. Det måste dom göra varje morgon och varje... **kväll (50)**. Sedan kryper dom ner i sängen. Mamma läser en godnattsaga för dom och tackar för den goda tårtan. Sally och Josef är nöjda och längtar redan till nästa dag. Då har pappa nämligen lovat att dom ska få äta tårta till frukost.

Slut!

Bilaga 4 - Bildstöd

Bilderna representerar målorden *källare*, *läskig*, *rund*, *slänga* och *skog*.



Bilaga 5 - Informationsbrev till vårdnadshavare

Vi heter Annie och Anna och är logopedstudenter vid Lunds universitet. Vi går sista terminen och skriver just nu vår magisteruppsats. Syftet med uppsatsen är att samla in ord som ofta förekommer i barnböcker för åldrarna 3–6 år. Sedan ska vi plocka ut de vanligaste förekommande orden och skapa ett ordförrådstest för att undersöka hur väl barn i de här åldrarna behärskar orden. Vi har utformat ett mindre ordförrådstest som avser att undersöka detta. Barnen som deltar ska vara enspråkiga och inte ha några problem gällande hörsel eller språk.

Genom att samla in ord som ofta förekommer i barnböcker kan vi identifiera vilka semantiska områden som barn i detta åldersspann förväntas behärska. Med dessa insikter kan vi få ökad kunskap om barns språkutveckling vad gäller ordförråd och ordbetydelse. Detta kan i sin tur ge värdefull information för bedömning och behandling av barn med språkstörning.

Testet är utformat som en saga och består av ett antal bilder som ditt barn ska benämna. Testningen utförs av oss och äger rum på en plats som vi kommer överens om. Barnets medverkan är frivillig och om barnet under testningen inte vill fortsätta avbryts den. Medverkan är pseudonymiserad. Det innebär att vid bearbetning av materialet kommer barnet få ett kodnamn och kodnyckel finns bara i ett exemplar och bevaras inlåst. Informationen som testningen ger kommer att delges till personer som är involverade i projektet samt i utbildningssyfte. Ni har rätt till att avbryta er medverkan när som helst. Om ni har några frågor är ni välkomna att kontakta oss via mail och telefon.

Med vänliga hälsningar,

Annie Karlsson
Logopedstudent
Telefonnummer: 0725242625
Mail: annie.karlsson.075@student.lu.se

Anna Svensson
Logopedstudent
Telefonnummer: 0768865476
Mail: log15asv@student.lu.se

Ketty Andersson
Handledare, Med. dr.
Telefonnummer: 046 17 78 33
Mail: ketty.andersson@med.lu.se

Kristina Hansson
Handledare, docent
Telefonnummer: 046 17 27 11
Mail: kristina.hansson@med.lu.se

Avdelningen för logopedi, foniatri och audiologi, Institutionen för kliniska vetenskaper, Lund, Lunds universitet, Universitetssjukhuset, 221 85 LUND.

Bilaga 6 - Medgivandeblankett

Vi har läst igenom informationsbrevet och förstått syftet med undersökningen. Vi vet att vi när som helst kan avbryta vår medverkan och vi tillåter vårt barn att delta i undersökningen.

Vårt barn har svenska som modersmål.

JA NEJ

Vårt barns språkutveckling är i nivå med jämnårigas.

JA NEJ

Vårt barn har haft logopedkontakt.

JA NEJ

Vårt barn har normal hörsel

JA NEJ Vet ej

Vårdnadshavare namn.....

Vårdnadshavare namn.....

Telefonnummer.....

Barnets förnamn.....

Barnets personnummer (6 siffror)

Vårdnadshavare namnteckning. Ort och datum

.....

Vårdnadshavares namnteckning. Ort och datum

.....

Undersökarens namn. Ort och datum

.....