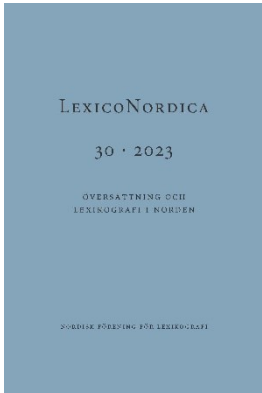


LexicoNordica

Titel:	Tvåspråkiga lexikala svensk-franska resurser för franska som främmande språk	
Forfatter:	Mårten Ramnäs	
Kilde:	LexicoNordica 30, 2023, s. 95-115	
URL:	https://tidsskrift.dk/lexn/issue/archive	

© 2023 LexicoNordica och författarna

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Tvåspråkiga lexikala svensk-franska resurser för franska som främmande språk

Mårten Ramnäs

This article provides an overview of some bilingual lexical resources for Swedish foreign language learners of French. The perspective is partly historical and includes dictionaries, parallel corpora and word lists. The overview shows that no major traditional bilingual Swedish-French dictionaries have been published since the 1990s. While many people use free online lexical services based on big data, the overview identifies some significant shortcomings in their current design. The article also explores the advantages and needs of a bilingual French word family list. The creation of a word family list for French would allow for a more systematic approach to vocabulary teaching in French at all levels in the Swedish educational system.

1. Inledning

Det finns ett tydligt samband mellan ordförrådets storlek hos inlärare av ett främmande språk och flera viktiga färdigheter som exempelvis läsförståelse och muntlig produktion (Schmitt, Jiang & Grabe 2011, Stæhr 2008). Språkinlärare på alla nivåer har därför all anledning att försöka utvidga sitt ordförråd. För inlärare på nybörjar- och mellannivå är olika typer av tvåspråkiga lexikala resurser ett viktigt hjälpmedel, medan avancerade inlärare i större utsträckning kan dra nytta av enspråkigt material (Svensén 2004:545–548). På gymnasiet läser 95 % av eleverna i franska inte längre än till steg 3 eller steg 4, trots att det finns möjlighet att läsa 7 steg (Skolverket 2022). Inom högre utbildning finner vi det absolut största antalet studenter på nybörjar- och grundkursnivå (Alberius m.fl. 2017:15, 77). Tillgången till adekvata tvåspråkiga

resurser av hög kvalitet är alltså av stor vikt för många inlärare, men även vid all typ av översättningsverksamhet. I föreliggande artikel kommer jag att granska några av de tvåspråkiga lexikala resurser som finns tillgängliga för svenska inlärare av franska samt reflektera över utvecklingen av dem fram till nu. I avsnitt 2 beskrivs med en kort historik följande resurser: ordböcker (2.1), parallellkorpora (2.2) och ordlistor (2.3). I avsnitt 2.3 blickar jag också framåt och pekar på behovet av en fransk frekvensordlista organiserad i ordfamiljer och försedd med svenska ekvivalenter. Avsnitt 3 avslutar artikeln med en kort summering.

2. Tvåspråkiga lexikala resurser

2.1. Ordböcker

I det första numret av *LexicoNordica* från 1994 skrev Martin Gellerstam att tvåspråkiga ordböcker för främmandespråksinlärare stod vid en skiljeväg där man ”får ta ställning till olika sätt att presentera, distribuera och uppdatera lexikalisk information” (Gellerstam 1994:51). Han såg en möjlig utveckling mot kontrastiva lexikon med fördjupad information, men förutspådde även att tvåspråkiga parallellkorpora skulle kunna bli värdefulla. Knappt trettio år senare kan vi konstatera att utvecklingen av moderna tvåspråkiga ordböcker till viss del har följt den bana som Gellerstam så framsynt förutspådde. De är oftast digitala, erbjuder fler sökmöjligheter och kan lättare uppdateras. De stora och omvälvande förändringar som Gellerstam i likhet med många andra förutspådde har emellertid låtit vänta på sig, åtminstone vad gäller franska ordböcker för svenskar. Låt oss först börja med en kort tillbakablick.

Under en stor del av 1900-talet dominerades marknaden i Sverige av två franska ordböcker: Thekla Hammars *Svensk-fransk ord-*

bok och Johan Visings *Fransk-svensk ordbok*. De första upplagorna av ordböckerna kom ut 1936 på Bonniers och Norstedts samägda läromedels- och ordboksförlag, Svenska bokförlaget. De tredje och sista upplagorna kom ut på 1950-talet, men nytryckningen skulle fortsätta ända fram till 1993 (elfte tryckningen), vilket visar hur lite som hände på ordboksmarknaden vad gäller franskan under det gångna seklet. Både Hammars och Visings ordböcker digitaliserades 2011 inom ramen för Projekt Runeberg. Även om de båda ordböckerna numera är föråldrade, kan de digitala versionerna fortfarande vara till gagn för inlärare som konfronteras med äldre texter. Sannolikheten att finna ekvivalenter till tidstypiska ord i de klassiska texterna från till exempel den franska realismen är ofta större än vid en sökning i nyare ordböcker.

Det skulle dröja fram till slutet av förra seklet innan nya stora ordboksprojekt såg dagens ljus. För svensk ordboksutgivning kom 1990-talet att bli en kort guldålder (Sköldberg & Mattsson 2016:112). Kanske innebar Sveriges inträde i EU att man såg ett större behov av nya resurser. Samma år som Sverige gick med i EU (1995) kom Enwall och Lötmarkers *Fransk-svensk ordbok* ut hos Natur & Kultur och tre år senare Norstedts franska lexikonpaket med *Norstedts stora fransk-svenska ordbok* och en svensk-fransk dito. Båda finns recenserade i *LexicoNordica*, den förra av Danell (1996), den senare av Bornäs (2000). Hur ser då läget ut 2023, ett kvarts sekel efter det att de två sist nämnda ordböckerna kom ut? Svaret är nedslående: Inga nya ordboksprojekt har startats sedan 1990-talet. *Fransk-svensk ordbok* (Natur & Kultur) trycktes bara i en upplaga, men kan sedan 2018 i elektroniskt format konsulteras genom WordFinders abonnemangsbaserade lexikontjänster. Norstedts franska ordböcker övergick till Nationalencyklopedin (NE) 2015 när det senare förlaget köpte upp Norstedts ordboksverksamhet. Några år senare (2018) kom den ut i en ny, lätt uppdaterad upplaga, som finns i både tryckt och elektronisk version. Med undantag för en smärre uppdatering av Norstedts ordbok är sålunda

det enda som hänt på storordboksfronten det senaste kvartsseket vad gäller franskan att de 25 år gamla ordböckerna har digitaliserats. Till saken hör att Bornäs (2000:236) redan för mer än tjugo år sedan konstaterade att ”[o]m man betraktar den svenska lexikonhistorien på franskans område frapperas man kanske mest av hur lite de nytkomna verken på grundläggande punkter skiljer sig från de äldre”. I sin recension gav Bornäs även tydligt uttryck för de förhoppningar på en snabb utveckling som fanns vid denna tid, i synnerhet för produktionsordböckerna. Det som Bornäs, Gellerstam och många andra i slutet av förra seklet inte kunde förutse var internet, som på kort tid nästan helt har slagit undan benen för nya kommersiella ordboksprojekt.

Som vi konstaterade är de gamla ordböckerna från 1990-talet trots allt digitaliserade och fortfarande flitigt använda, särskilt Norstedts (nu NE). Vad erbjuder dessa då i elektroniskt format som de tryckta versionerna inte gjorde? Att kunna höra orden uttalas och att texten är mindre kondenserad (förkortningar och plattållarsymboler har till stor del försvunnit) är exempel på förändringar. Utöver dessa förändringar torde emellertid fler sökmöjligheter vara den enskilt största skillnaden.

NE medger sökningar i språkproven (i NE kallat fraser) på källspråksidan. En sökning på *sale* 'smutsig' i NE ger sålunda träff även på fraser som står under andra ingångslemman än sökordet, till exempel *tu as une sale mine* 'du ser förskräcklig ut', som står under *mine*. I en elektronisk ordbok kan emellertid i princip all text göras sökbar. Att målspråksenheter, som tidigare var åtkomliga endast indirekt via källspråket och helt motiverades av detta, nu kan bli sökbara, väcker många frågor, särskilt när gamla ordböcker digitaliseras, vilka egentligen inte har utformats med dessa sökmöjligheter i åtanke (se Hannesdóttir 2013).

I WordFinder, som har införlivat Natur & Kulturs ordbok i sin tjänst, finns funktionen Fulltextsök, som ger träff i alla delar av ordboken, inte bara på källspråkssidan utan även på målspråkssi-

dan. I en från början enkelriktad fransk-svensk ordbok (en receptionsordbok för svenskar) kan man alltså söka på de svenska ekvivalenterna, och, om uttrycket tillåts, gå bakvägen till ett franskt ingångslemma. En receptionsordbok för svenskar kan därmed i princip tjäna som produktionsordbok för fransmän, men då måste vi komma ihåg att de svenska ekvivalenterna i en sådan ordbok helt motiveras av sina respektive franska lemmen, medan dessa senare är resultatet av en metodisk selektion. Som produktionsordbok för fransmän har den alltså uppenbara brister. Att tillåta sökningar av sådan typ i dubbelriktade ordböcker skulle ändå kunna ge ett mervärde. Om vi tar ordet *smutsig* som exempel så nöjer sig en vanlig svensk-fransk ordbok ofta med att ge *sale* 'smutsig' och *malpropre* 'oren' som ekvivalenter. En sökning på *smutsig* i WordFinder (Natur och Kulturs fransk-svenska ordbok) kompletterar med till exempel *tâché* 'fläckig', *terreux* 'jordig' eller *salingue* (vardagligt uttryck för smutsig), som alltså samtliga har *smutsig* som en av sina ekvivalenter. Även ekvivalenter av förklarande typ som innehåller sökordet ger träff (till exempel *salissant* 'som lätt blir smutsig'). "Bakvägssökningar" ger på så sätt möjlighet att upptäcka fler ekvivalenter, även om dessa ofta är mer perifera och kontext- eller registerberoende. Som vi konstaterade ovan ger Fulltextsök träff i alla delar av ordboken, även i den text som har använts för ekvivalentdifferentiering (glossning, kontextangivelser, etc.). En sökning på *ljud* ger träff på ingångslemmat *mordant* eftersom det följs av kontextangivelsen "om ljud 'genomträngande'". Genom att ge åtkomst till uppslagsordens innehållssida (ekvivalenterna) kan Fulltextsök sägas öppna för det onomasiologiska perspektivet och ge fragmentariska svar på frågor som exempelvis: Vilka ord används för begreppet 'ljud' i franskan? Sökträffar av detta slag blir emellertid lätt förvirrande och det gäller att användaren kan tolka dem på rätt sätt. Det kan vara anledningen till att NE, som riktar sig till en bred publik, bland annat skolelever, har valt att begränsa sökmöjligheterna. Sökfunktionerna i båda ordböckerna är, när allt

kommer omkring, tämligen enkla och det bör återigen påpekas att ordböckerna inte från början utarbetats för sökmöjligheter som dessa – det är uppenbart att det i grunden rör sig om digitaliserade pappersordböcker (Svensén 2004:518).

Vi kan alltså konstatera att det inte kommit några nya stora ordboksprojekt vad gäller franskan sedan 1990-talet. Norstedts ordbok blev uppköpt av NE, som är det enda kvarvarande förlaget med ordboksverksamhet i traditionell mening. NE:s ordbok uppdateras fortfarande, men år 2018 bestod redaktionen av endast två personer med ansvar för ett tjugotal dubbelriktade ordböcker (Almegård 2018:48). Utbudet av gratisjänster på nätet har gjort det svårt för kommersiella ordboksutgivare. Josephson (2018:251) menar att statlig finansiering i någon form är nödvändig. Även Sköldberg & Mattsson (2016:122–123) ansluter till denna linje och framhåver det faktum att tvåspråkiga ordböcker till hälften består av det svenska språket. Risken är, menar författarna, att vi via gratisprodukter möter en svenska som inte är så standardiserad som tidigare.

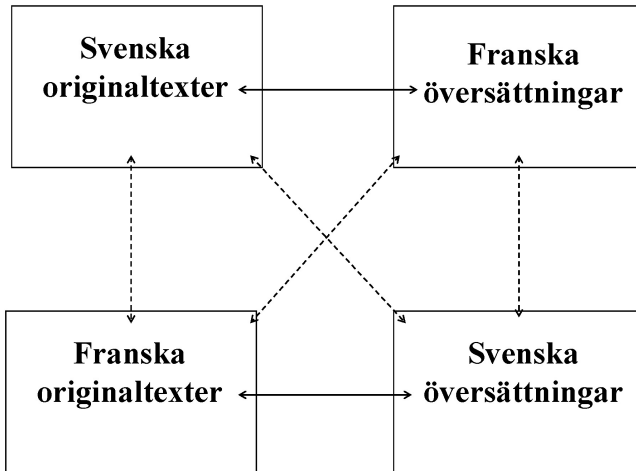
Att ordböckerna, de som överlevt, inte har utvecklats mer än de har gjort kan förutom marknadsläget också bero på att svårigheterna med att framställa all den information som det digitala formatet potentiellt tillåter, på ett sätt som är konsekvent och inte alltför komplicerat, har visat sig vara större än vad man tidigare trodde, i synnerhet när ordböcker också ska kunna konsulteras på mobiltelefoner.

2.2. Parallellkorpusar och deras användning i tvåspråkiga lexikala resurser på nätet

Som vi konstaterade under 2.1 förutspådde Gellerstam (1994) att tvåspråkiga parallellkorpusar skulle kunna utgöra en viktig lexikal resurs i framtiden och på den punkten ser han ut att ha fått rätt. Från att i första hand ha varit ett hjälpmedel för lexikografer,

översättare och språkforskare kan idag alla internetanvändare genom några klick få fram konkordanser från mycket omfattande korpusar.

Under början av 2000-talet skapades flera svensk-franska parallellkorpusar (se till exempel Kortteinen & Ramnäs 2002, Mossberg 2006, Wastesson 2010) enligt den modell som gällde för den engelsk-svenska parallellkorpusen, ENPC (Aijmer, Altenberg & Johansson 1996). Parallellkorpusar uppbyggda på detta sätt har många fördelar eftersom de medger jämförelser av flera olika slag.



Figur 1: Modell för dubbelriktade parallellkorpusar.

Originaltexter på svenska och franska kan jämföras med sina respektive översättningar. En viktig poäng är att jämförelsen kan göras i båda riktningarna (se heldragna dubbelriktade pilar i figur 1). Man kan gå från källtext till måltext som i exempel 1 och 3 nedan och från måltexten tillbaka till källtexten, som i exempel 2 och 4 (se Ramnäs 2008). Eftersom språket i översatta texter tenderar att påverkas av källspråket, så kallad translationese (Gellerstam 1986), kan en korpus bestående av original och översättningar på

båda språken bidra till att ge en mer nyanserad bild av olika lexikala ekvivalensförhållanden mellan två språk. I fallet med svenskans modala *fa* och franskans *faire* + infinitiv är uttrycket *avoir le droit de* 'ha rätt att' (1) och kausativa analytiska konstruktioner som *fa* + infinitiv (3) mycket vanligare som översättningsmotsvarigheter i riktningen källspråk till målspråk. Genom att undersöka riktningen målspråk till källspråk visar sig andra uttryck vara vanligare (2 & 4). I en korpus med originaltexter på bara ett av språken blir bilden alltså inte lika komplett.

- | | | | |
|-----|--|---|---|
| (1) | Du <u>får inte</u> | → | Tu <u>n'as pas le droit</u> |
| (2) | Du <u>får inte</u> | ← | Tu <u>ne dois pas</u> |
| (3) | Son regard immobile, inexpressif, <u>me fait penser au</u> regard des fauves | → | Hans orörliga, uttryckslösa blick <u>får mig att tänka på</u> ett vilddjurs blick |
| (4) | Parfois, <u>ils lui faisaient penser à</u> une troupe d'oiseaux multicolores | ← | Ibland <u>påminde de henne om</u> en flock färggranna fåglar |

I korpusar uppbyggda enligt figur 1 kan även översatta texter jämföras med original på samma språk för att bättre uppmärksamma källspråkspåverkan i form av över- och underanvändning av ord, uttryck och strukturer i översatta texter (streckade diagonala pilar i figur 1). Slutligen kan originaltexter på båda språken jämföras med varandra (streckade lodräta pilar). Parallellkorpusar av ovan beskrivet slag innehåller alltså fyra delkorpusar och kan på en och samma gång sägas vara dubbelriktade översättningskorpusar (samma texter på två olika språk) och jämförbara korpusar (liknande texter på olika språk). De svensk-franska parallellkorpusarna som skapades på 2000-talet har legat till grund för flera avhandlingsprojekt som behandlar olika problem inom kontrastiv lexikologi. Tyvärr är tillgången till korpusarna begränsad till forskning. Frågor om upphovsrätt bör alltså redas ut innan en

korpus etableras, om tanken är att den ska bli helt öppen för allmänheten. En komplikation i fallet med parallellkorpusar är att tillstånd måste inhämtas för både originaltext och översättning. Trots svårigheterna med upphovsrätten skulle det vara av intresse att utreda om äldre korpusar, uppbyggda enligt principen som beskrivits ovan, kan göras tillgängliga på något sätt, kanske genom att begränsa kontexten och göra ordningen på sökträffarna slumpmässig så att texterna i sin helhet aldrig kan kopieras. De korpusar som finns fritt tillgängliga på nätet är nämligen behäftade med en rad problem, vilket vi kommer att se i det som följer.

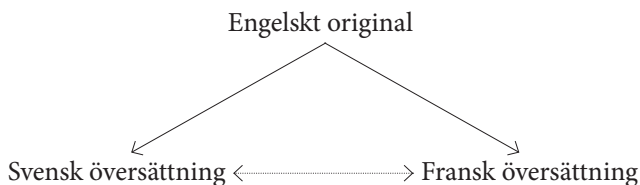
Vi har nu berört dubbelriktade svensk-franska parallellkorpusar uppbyggda på ett särskilt sätt, och som i de flesta fall endast finns tillgängliga för forskare. På nätet finns emellertid fritt tillgängligt ett stort antal flerspråkiga korpusar där både svenska och franska ingår. EU har varit en föregångare på området genom att göra många korpusar öppna¹. EU finansierar även jätteprojektet ParaCrawl (Bañón m.fl. 2020) som samlar in flerspråkiga texter på nätet. För att få en representation av talat språk används ofta undertexter till filmer. Opensubtitles2018 (se även Lison & Tiedemann 2016) är exempel på en korpus bestående av användargenererade undertexter till filmer på ett sextiotal olika språk. Stora korpusar är viktiga för att exempelvis träna maskinöversättningsprogram. För en privatperson utan språkteknologisk kompetens är de två sistnämnda korpusarna svåra att använda direkt, men indirekt, via företag som erbjuder språktjänster på nätet, kommer korpusarna privatpersoner till godo.

Reverso är ett franskt företag, som bland annat erbjuder gratis ordbokstjänster och som kan tjäna som exempel för många andra liknande företag (till exempel Glosbe). När man söker på ett ord i Reverso får man dels upp de vanligaste översättningsekvivalenterna och frekvensuppgifter för dessa, dels konkordanser från de

1 På följande sida listas flera av korpusarna: <joint-research-centre.ec.europa.eu/language-technology-resources_en>

stora korpusarna. Eftersom ingen ekvivalentdifferentierande text erbjuds tjänar konkordanserna som vägledning för hur ordet kan användas. Det faktum att man kan söka på flera ord är till stor hjälp eftersom kontexten då kan specificeras. Så här långt ser konceptet ut att vara mycket användbart och visar att bruket av parallellkorpusar har ökat och spritts. Eftersom ordbokstjänsten nästan helt bygger på korpusdata finns emellertid också all anledning att peka på några av de problem de stora korpusarna behäftas med. Paracrawl går, som vi tidigare nämnt, ut på att med datorers hjälp söka efter flerspråkiga websidor på internet. På statistisk väg och med hjälp av maskinöversättning bedöms sannolikheten för att två eller flera texter på olika språk har samma innehåll. Texter som uppfyller fastställda kriterier länkas sedan på meningsnivå. När man tittar närmare på parallellkonkordanser från ParaCrawl är det emellertid uppenbart att de innehåller stora mängder maskinöversatt text eller översättningar av låg kvalitet. Till detta kommer problemet med att originaltexterna oftast varken är svenska eller franska.

Det som i konkordanserna ser ut att representera resultatet av en översättning mellan svenska och franska (streckade pilar i figur 2) är i själva verket ofta två separata översättningar från ett helt annat språk, vanligen engelska (heldragna pilar). Det tredje språket fungerar då som brygga.



Figur 2: Engelska som brygga i en svensk-fransk parallellkorpus.

Opensubtitels2018 har inte samma problem med maskinöversatta texter, men korpusen delar problemet med bristen på svenska

originaltexter, och även i viss mån franska originaltexter. Om ett av de två språken är underrepresenterat blir jämförelserna enkelriktade och det andra språket präglar bilden av olika ekvivalensförhållanden. Om originalspråket är engelska kommer det språket att präglade bilden. I 2.1 konstaterades att målspråksenheter i en ordbok bestäms av källspråkets lemmalista. När en ordbokstjänst huvudsakligen baserar sig på korpusar med originaltexter på ett annat språk än svenska uppstår liknande problematik – svenska ingångslemman finns där för att de är resultatet av en översättning till svenska, inte för att de är resultatet av en metodisk selektionsprocess eller för att de har belagts i texter som ursprungligen är skrivna på svenska. Kulturspecifika svenska ord (*nyponsoppa*, *falukorv*, *friggebod*) tenderar till exempel att vara dåligt representerade. Generellt sett innebär givetvis också två översättningar större risk för betydelseförskjutningar än en.

Om man söker på det svenska verbet *samla* anger Reverso *collecter* som den första och vanligaste franska ekvivalenten. Detta franska verb är i själva verket relativt lågfrekvent² och kontextberoende i jämförelse med till exempel *réunir* och *rassembler* som ofta anges som ekvivalenter i traditionella ordböcker. Man kan misstänka att det rör sig om inflytande från engelska originaltexter (engelskans *collect*).

EU-texter håller i allmänhet jämnare och, relativt sett, högre kvalitet, men det är notoriskt svårt att veta vad som är originalspråket. Det torde under alla förhållanden mycket oftare vara engelska (eller möjligen franska) än svenska. Inom EU översätts texter vanligen med CAT-verktyg, vilket innebär att översättarna på olika sätt måste förhålla sig till tidigare översättningar. Normalt skall den så kallade punktregeln tillämpas, vilket innebär att en mening i källtexten ska svara mot en mening i måltexten. Texter-

2 I Lonsdale & Le Bras (2009) frekvensordlista tillhör verbet *collecter* inte franskans 5 000 vanligaste lemman medan *réunir* har rang 824 och *rassembler* rang 1635. I den svenska Kelly-ordlistan har *samla* rang 837.

na genomgår också olika typer av granskningsprocesser för ökad koherens mellan språken (Bendegard 2014). Det finns nu 552 språkkombinationer i EU, vilket innebär att indirekt översättning förekommer. De stora kraven på konsekvens och det faktum att översättningar till olika språk inte sällan utarbetas parallellt gör att man i många sammanhang pratar om olika språkversioner av en text, snarare än om översättningar. Inte sällan avviker texternas språk så mycket från vanlig svenska att man föredrar att betrakta dem som EU-texter på svenska (Bendegard 2014:179–182).

En ny typ av tjänst är de ordböcker till vilka användarna själva bidrar. Reverso har en sådan kallad ”dictionnaire collaboratif”. Den innehåller emellertid i nuläget många luckor och erbjuder ingen hjälp för ekvivalentdifferentiering.

Sammanfattningsvis kan vi konstatera att kvaliteten på de gratis ordbokstjänster som finns på nätet är skiftande. Om de i huvudsak förlitar sig på korpusdata måste den vara balanserad, representativ och kontrollerad. Även användargenererat innehåll bör granskas. Just nu är tendensen att samla stora mängder data, inte sällan till förfång för balans och kvalitet. Tjänsterna kommer förhoppningsvis att utvecklas och bli bättre. I nuläget kan de tjäna som ett bra komplement till traditionella ordböcker, men knappast ersätta dem.

2.3. Ordlistor

Nya ord kan läras in direkt, via explicita metoder vars uttalade syfte, åtminstone delvis, är att inläraren ska införliva nya ord i sin vokabulär, eller indirekt (implicit) i upprepade möten med skrivet eller talat språk i situationer (till exempel läsning eller radiolyssnande) där syftet inte är att lära in nya ord. I en främmandespråkskontext där undervisningstiden i många fall är begränsad till två-tre timmar i veckan är de flesta forskare överens om att explicita metoder är nödvändiga (Laufer 2017).

För att avgöra vilken vokabulär som är lämplig för en given nivå eller kurs, för att utarbeta övningar och för att mäta och bedöma inlärares ordförråd har ordlistor i alla tider upprättats. Den begynnande inlärares har exempelvis all anledning att börja med att lära in de mest frekventa och användbara orden. En ordlista skall i sig inte ses som en metod utan sätter ett mål för undervisningen och tjänar som underlag för övningar och tester. Att en ordlista förses med ekvivalenter på inlärares modersmål har ett stort värde.

There are numerous ways of conveying the meaning of an unknown word [...]. However, studies comparing the effectiveness of various methods for learning always come up with the result that an L1 translation is the most effective (Nation 2003:3–4).

I Sverige användes under lång tid *Centrala ordförrådet i franska* (Schlyter 1951). Den omfattade 2 200 franska lemman ordnade i nästalfabetisk listuppställning med svenska ekvivalenter. Senare kom en ny version kallad *Svensk-franska centrala ordförrådet* (Schlyter 1988), med cirka 4 000 svenska lemman ordnade i striktalfabetisk listform och försedda med franska motsvarigheter.

Sedan 2008 har samtliga svenska lärosäten där man kan läsa franska enats om en ordlista, den s.k. riksprovslistan (fortsättningsvis RL), som ska gälla för grundkursen (se även Lindqvist & Ramnäs 2022). Den omfattar knappt 4 000 lemman, också de ordnade i en striktalfabetisk lista. Orden testas på det så kallade Riksprövet i fransk grammatik och vokabulär. Den officiella versionen av RL är enspråkig, men de flesta lärosäten har efterhand försett den med ekvivalenter. Det förenklar inte bara inläringen för studenterna utan förtydligar också *vad* de förväntas kunna om orden (hörfrekventa ord tenderar att vara polysema). Att listan nu i de flesta fall är tvåspråkig är positivt, men att olika tvåspråkiga versioner av listan förekommer parallellt är olyckligt. För hög kva-

litet bör man sträva efter att en officiell version utarbetas med stöd av lexikografisk kompetens.

Som lexikografisk produkt har tvåspråkiga ordlistor inte rönt samma intresse som ordböcker trots den stora betydelse de har för många studenter. RL är en striktalfabetisk lista som vi konstaterade ovan. I en ordbok underlättar alfabetisk ordning snabb åtkomst, men bör det verkligen vara ett tungt vägande kriterium för en ordlista avsedd för språkundervisning? Rey (1977:20) menar att den alfabetiska ordningen på många sätt kan sägas vara absurd. Ord kan ordnas på andra sätt, exempelvis tematiskt efter betydelse. Att ordna morfologiskt besläktade ord i nästen som bryter den alfabetiska ordningen utgör ytterligare ett exempel på alternativ makrostruktur. Svensén (2004:443–444) hävdar att syftet med en sådan struktur i regel är att underlätta för ordförståelse och ordinläring. Svensén utvecklar inte detta påstående, men den bakomliggande idén torde vara att en sådan struktur gynnar inlärarens morfologiska medvetenhet. Kunskap om de vanligaste affixen gör att man kan fokusera på den gemensamma stammen i en ordfamilj istället för att lära in varje ord separat. Det kan därför tyckas märkligt att många moderna ordlistor är striktalfabetiska. Förklaringen kan vara att det är enkelt att generera alfabetiska frekvensordlistor utifrån korpusar. Att ordna dem i nästen är däremot mycket mer tidkrävande.

I en nästalfabetiskt ordnad ordbok eller ordlista bildar nästena ordfamiljer. Inom engelsk andraspråkforskning används just ordfamiljer, snarare än lemman, som enhet när man vill mäta det antal ord en inlärare kan. Bland annat har Bauer & Nation (1993) föreslagit en hierarkisk modell med sju nivåer som operationellt definierar begreppet ordfamilj (se figur 3). Första nivån utgörs av ordformer med samma stavning (*types*), den andra nivån av lemmats böjda former och därefter följer fyra nivåer som successivt introducerar ett nittiotal avledningsaffix. Den sjunde och sista nivån är något oklart definierad och används sällan.

Nivå	Ordfamiljen <i>develop</i>	Antal avledningsaffix samt exempel på affix
1		
2	<i>develop, develops, developed, developing</i>	
3	<i>developable, undevelopable, developer(s), undeveloped</i>	10: <i>-able, -er, -ish, -ness, -th, un-</i>
4	<i>development(s), developmental, developmentally</i>	11: <i>-al, -ation, -ism, -ment, -ize, in-</i>
5	<i>developmentwise, semideveloped, antidevelopment</i>	50: <i>-ance, -ery, -hood, anti-, semi-, -wise</i>
6	<i>redevelop, predevelopment</i>	12: <i>-ee, -ic, -ist, -ive, pre-, re-</i>

Figur 3: Bauer & Nations modell för ordfamiljer.

Bauer & Nation (1993) menar att ordfamiljer har potential att minska inlärningsbördan och öka inlärnarnas morfologiska medvetenhet. Utgångspunkten för Bauer & Nations (1993) modell är att en inlärare ska förstå eller enkelt kunna lära in andra medlemmar i ordfamiljen så fort en medlem har lärts in. Antalet affix kan tyckas vara stort, men eftersom modellen är hierarkisk kan successivt fler avledningsaffix läggas till för progression i undervisningen. Störst nytta har inlärare av ordfamiljer när det gäller receptiva kunskaper. Morfologisk medvetenhet hjälper inlärare att dra slutsatser om betydelsen av okända ord. När det gäller produktiva kunskaper kan inlärare däremot inte alltid veta vilka avledningar som har lexikaliserats.

Ramnäs & Lindqvist (2021) kom fram till att Bauer & Nations (1993) hierarkiska modell för ordfamiljer i princip är överförbar även till franskan. Det finns några enstaka enspråkiga franska ordböcker baserade på ordfamiljer, men de för oftast ihop ord i mycket stora familjer där sambanden mellan orden långt ifrån alltid

är transparenta för en vanlig språkanvändare. Äldre tvåspråkiga ordböcker var ofta nästalfabetiska, och det gällde även *Centrale ordförrådet i franska* (Schlyter 1951). Problemet med dessa var att endast lemmen avledda med suffix ordnades i nästen, aldrig prefixavledda lemmen.

I en striktalfabetisk ordlista skulle orden *stable, instable, stabilité, instabilité, stabiliser, déstabiliser, stabilisation, déstabilisation, stabilisateur, déstabilisateur* etc. förekomma på olika ställen. En del av orden i familjen är dessutom lågfrekventa och platsar därmed inte i frekvenslistor av RL:s typ där lemmaselektionen i huvudsak baseras på frekvens. Fördelen med ordfamiljsbegreppet är att det skulle skilja på den sorts lågfrekventa lemmen som inte har andra högfrekventa medlemmar i sin familj (till exempel *meugler* 'råma') och sådana lemmen som är lika lågfrekventa (*déstabilisateur* 'destabiliserande'), men som har andra högfrekventa medlemmar i sin familj och därmed lättare kan förstås eller läras in.

Med utgångspunkt från en indelning i ordfamiljer uppenbarar sig i nuvarande version av RL många inkonsekvenser. Både *lutter* 'kämpa' och *lutte* 'kamp' finns med, men endast *oublier* 'glömma', inte *oublir* 'glömska'. På samma sätt finns *inonder* 'svämma över' och *inondation* 'översvämning' med, men bara *interprétation* 'tolkning', inte *interpréter* 'tolka'.

I en frekvensordlista på 5 000 lemmen visade Ramnäs & Lindqvist (2021) att man genom att gruppera alla regelbundet avledda substantiv med *-ation* (*annuler – annulation*), substantiv och adverb avledda med *-ment* (*encourager* 'uppmuntra' – *encouragement* 'uppmuntran'; *rapide* 'snabb' – *rapidement* 'snabbt'), substantiv bildade ur verbstammar (*rêver* 'drömma' – *rêve* 'dröm'), samt infinitiv- och participformer (*blessé* 'skadad', *gagnant* 'vinnare', *rire* 'skratt') kunde gruppera cirka 25 % av samtliga lemmen. Utöver detta kan en stor mängd lågfrekventa avledningarna flyttas upp och integreras i listan. Detta är exempel på affix som skulle kunna ut-

göra en första nivå i en hierarkisk modell för franskan. Fler nivåer kan sedan byggas på. Även fler betydelser för orden kan läggas till. I digitalt format skulle ordlistan på så sätt kunna anpassas efter nivån.

3. Några avslutande ord

I denna artikel har jag med en kort historik beskrivit ett antal kontrastiva lexikala resurser som finns tillgängliga för svenska inlärare av franska som främmande språk. Genomgången visar att utvecklingen av förlagsutgivna tvåspråkiga svensk-franska resurser i det närmaste helt avstannat. Utvecklingen vad gäller andra tvåspråkiga (gratis) resurser på nätet har däremot gått fort. Sådana resurser kan fungera som ett bra komplement till de mer traditionella, men i nuläget kan de knappast ersätta dem. Att utveckla de traditionella resurserna kommer i framtiden troligtvis att kräva statlig finansiering i någon form.

Många undersökningar visar att de som har störst nytta av tvåspråkiga lexikala resurser är nybörjare och inlärare på mellannivå, inte de mest avancerade inlärarna, som har större möjlighet att använda enspråkiga resurser (Svensén 2004:545–548). Kanske är det också på de lägre nivåerna, där det stora flertalet inlärare befinner sig, som man borde satsa mer så att de möts av material som håller hög kvalitet. En tvåspråkig familjebaserad fransk ordlista, framtagen med stöd av lexikografisk expertis, där storleken på familjerna och antalet ekvivalenter kan anpassas efter nivå skulle vara till stort gagn för svensk utbildning, både inom ungdomsskolan och på universiteten. Med vetenskapligt framtagna tvåspråkiga lexikala resurser av hög kvalitet blir det lättare för elever och studenter att tillägna sig den grundläggande vokabulär som är nödvändig för fortsatta studier på högre nivåer.

Litteratur

Ordböcker

Enwall, Jeannie & Rut Lötmarker (1995): *Fransk-svensk ordbok*. Stockholm: Natur och Kultur.

Glosbe = <sv.glosbe.com> (mars 2023).

Hammar, Thekla (1936): *Svensk-fransk ordbok*. Stockholm: Svenska bokförlaget/Norstedts. (Även tillgänglig på <runeberg.org/authors/hammathe.html > (augusti 2023).)

Kelly = <spraakbanken.gu.se/resurser/kelly> (juni 2023).

Lonsdale, Deryle & Yvon Le Bras (2009): *A Frequency Dictionary of French*. New York: Routledge.

Nationalencyklopedins ordböcker = <ne.se/ordböcker> (mars 2023).

Norstedts stora fransk-svenska ordbok (1998). Stockholm: Norstedts.

Projekt Runeberg = <runeberg.org> (juli 2023).

Reverso = <dictionary.reverso.net/> (mars 2023).

Schlyter, Börje (1979 [1951]): *Centrala ordförrådet i franskan*. Stockholm: Almqvist & Wiksell läromedel.

Schlyter, Börje (1988): *Svensk-franska centrala ordförrådet*. Lund: Dialogos.

Vising, Johan (1936): *Fransk-svensk ordbok*. Stockholm: Svenska bokförlaget/Bonniers. (Även tillgänglig på <runeberg.org/authors/visinper.html > (augusti 2023).)

WordFinder = <app.wordfinder.com/sv> (mars 2023).

Annan litteratur

Alberius, Lars, Markella Callin, Angela Hoffman, Sverker Johansson & Coco Norén (2017): *Behovet av en språkstrategi för Sverige: Slutrapport med förslag från Arbetsgruppen för språkämnen*. Rapport från SUHF. Dnr. 0051-17. (14/075.)

- Aijmer, Karin, Bengt Altenberg & Mats Johansson (1996): Text-based contrastive studies in English. Presentation of a project. I: Karin Aijmer, Bengt Altenberg & Mats Johansson (red.): *Languages in contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*. Lund: Lund University Press, 73–85.
- Almegård, Mats (2018): Vem ska rädda ordboken? I: *Språktidningen* 6, 46–52.
- Bañón, Marta, Pinzhen Chen, Barry Haddow, Kenneth Heafield, Hieu Hoang, Miquel Esplà-Gomis, Mikel L. Forcada, Amir Kamran, Faheem Kerfi, Philipp Koehn, Sergio Ortiz Rojas, Leopoldo Pla Sempere, Gema Ramírez-Sánchez, Elsa Sarrías, Marek Strelec, Brian Thompson, William Waites, Dion Wiggins & Jaime Zaragoza (2020): ParaCrawl: Web-Scale Acquisition of Parallel Corpora. I: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4555–4567, Association for Computational Linguistics.
- Bauer, Laurie & Paul Nation (1993): Word families. I: *International Journal of Lexicography*, 6 (4), 253–279.
- Bendegård, Saga (2014): *Begriplig EU-svenska? Klarspråksarbetets förutsättningar inom den interinstitutionella översättningsprocessen*. Uppsala: Uppsala universitet.
- Bornäs, Göran (2000): Ett franskt lexikonpaket från Norstedts. Recension av: *Norstedts stora fransk-svenska ordbok. Le Grand Dictionnaire français-suédois de Norstedts*. Stockholm: Norstedts Ordbok. I: *LexicoNordica* 7, 223–238.
- Danell, Karl Johan (1996): Recension. Jeannie Enwall och Ruth Lötmarker. 1995. *Fransk-svensk ordbok*. Stockholm: Natur och Kultur. I: *LexicoNordica* 3, 223–229.
- Gellerstam, Martin (1986): Translationese in Swedish novels translated from English. I: Lars Wollin & Hans Lindquist (red.): *Translation studies in Scandinavia*, Lund studies in English 75, 88–95. Malmö: Liber/Gleerup.
- Gellerstam, Martin (1994): Produktionsordböcker – vad är det? I: *LexicoNordica* 1, 43–52.

- Hannesdóttir, Anna Helga (2013): Lemman och ekvivalenter i nya roller – en reviderad ordbokstypologi. I: Ruth Vatvedt Fjeld & Marit Hovdenak (red.): *Nordiske Studier i Leksikografi* 12. Oslo: Nordisk forening for leksikografi, 193–211.
- Josephson, Olle (2018): *Språkpolitik*. Stockholm: Morfem.
- Kortteinen, Pauli & Mårten Ramnäs (2002): Om en svensk-fransk parallellkorpus – Nytt och tillämpningar. I: *Humaniora – en akademisk fråga. Humanistdag-boken* nr 15, 141–154. Göteborg: Göteborgs universitet.
- Laufer, Batia (2017): The Three ‘I’s of second language vocabulary learning: Input, instruction, involvement. I: Eli Hinkel (ed.): *Handbook of research in second language teaching and learning*, vol. III, 343–354. London: Routledge.
- Lindqvist, Christina & Mårten Ramnäs (2022): Élaboration d’une liste de vocabulaire dans un contexte FLE. I: *Studia romanica posnaniensia*, 49 (4), 65–74.
- Lison, Pierre & Jörg Tiedemann (2016): OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. I: *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, 923–929.
- Mossberg, Mari (2006): *La relation de concession. Étude contrastive de quelques connecteurs concessifs français et suédois*. Acta Wexionensia. Växjö: Växjö universitet.
- Nation, Paul (2003): The role of the first language in foreign language learning. I: *Asian EFL Journal*, 5 (2), 1–8.
- Ramnäs, Mårten (2008 [2006]): *Étude contrastive du verbe suédois få dans un corpus parallèle suédois–français* (ACTA Universitatis Gothoburgensis, Romanica Gothoburgensia 62). Göteborg: Göteborgs universitet.
- Ramnäs, Mårten & Christina Lindqvist (2021): Le concept de famille de mot appliqué au français: réflexions et simulations. I: *Les langues modernes*, 3, 45–53.

- Rey, Alain (1977): *Le lexique, images et modèles du dictionnaire à la lexicologie*. Paris: Colin.
- Schmitt Norbert, Xiangying Jiang & William Grabe (2011): The percentage of words known in a text and reading comprehension. I: *The Modern Language Journal*, 95(1), 26–43.
- Skolverket (2022): Skolutveckling – Statistik - Gymnasieskolan - Betyg och studieresultat - Riksnivå - Läsår 2021/22 - Tabell 12. <skolverket.se> (mars 2023).
- Sköldberg, Emma & Christian Mattsson (2016): Ful- och finlexikografi? Om ordboksverksamhet i Sverige i dag och i morgon: I: *Framtidens lexikografi. Rapport från ett symposium den 5 oktober 2012*. Meijerbergs arkiv för svensk ordforskning 42, 111–129.
- Stæhr, Lars Stenius (2008): Vocabulary size and the skills of listening, reading and writing. I: *The Language Learning Journal*, 36 (2), 139–152.
- Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedt.
- Wastesson, Sofia (2010): *L'atténuation en traduction. Etude traductologique des marqueurs de probabilité en français et en suédois*. Linnaeus University Dissertations, Nr 28/2010. Växjö, Kalmar: Linnaeus University Press.

Korpusar

- OpenSubtitles2018. <opus.nlpl.eu/OpenSubtitles-v2018.php> (mars 2023).
- ParaCrawl. <paracrawl.eu/> (mars 2023).

Mårten Ramnäs
 Universitetslektor
 Göteborgs universitet
 Institutionen för språk och litteraturer
 SE-405 30 Göteborg
 marten.ramnäs@sprak.gu.se