

Anglismit ja tekstin luettavuus

Jaakko Lehtonen ja Hannele Heikkinen (Jyväskylä)

1. Taustaa

Jyväskylän yliopistossa vuosina 1977—80 tehtyyn anglismitutkimukseen sisältyi osa, jossa n. 150 Helsingin, Lahden ja Kuopion ammattikoulujen ja teknisten oppilaitosten opiskelijaa joutui lukemaan ääneen kolme lyhyttä, tyylliltään ja sanastoltaan erilaista tekstiä: 1) ns. poptekstin, joka pyrki jäljittelemään koululaislangin ja poplehtien tyyliä ja jossa oli runsaasti nuorison kieleen kuuluvaa vierasperäistä sanastoa, 2) ns. normaalitekstin, joka oli katkelma erään pankin asiakkailleen osoittamasta tiedotteesta, sekä 3) ns. sivistyssanatekstin, joka oli asiatyylinen, runsaasti sivistyssanoja sisältävä, sisältönsä puolesta täysin tekaistu teksti. Näiden avulla haluttiin selvittää, miten tekstiin sisältyvät mukautumattomat vierassanat vaikuttavat tekstin luettavuuteen. Kukin suoritus transkriboitiin ja koehenkilöiden tekemät lukuvirheet, takeltelut sekä vierasperäisten sanojen /b d g f š/ -äänteitä tarkoittavien kirjainten ääntämistapa¹ kooditettiin niin, että koehenkilöiden ääntämissuorituksia voitiin verrata tutkimuksen taustatietomuuttujiin ja kirjallisen kokeen tehtäviin. Lisäksi mitattiin kunkin koehenkilön tekstin luennan kokonaisnopeus, taukojen prosenttiosuus kokonaisajasta sekä ääntönopeus (mittausmenetelmistä ks. Lehtonen, toim., 1978). Transkriptiosta laskettiin taukojen, takeltelujen ja lukuvirheiden määrät ja niiden jakautuminen tekstin eri kohtiin. Kaikkiaan lukukokeessa oli 83 kooditettua kohtaa, joiden yhteinen reliabiliteettikerroin (OSANA-osioanalyysi) oli 0,766. Koetta saattoi siis pitää varsin reliaabelina. Tulosten perusteella oli mahdollista selvittää, miten tekstin rakenne, sen tyyli ja koehenkilöiden sosiaalinen tausta vaikuttavat lukemisen sujuvuuteen. Samoin oli mahdollista saada tietoa siitä, miten lukemistestin tulokset korreloivat kirjallisen kieli- ja tyyliäidon kokeen tuloksien kanssa.

¹ Ääntämisseikkoja käsitellään yksityiskohtaisemmin Tampereella keväällä 1981 pidettyjen 10. fonetiikan päivien esitelmäjulkaisun artikkelissa (Heikkinen 1981).

Jaakko Lehtonen ja Hannele Heikkinen

2. Mitä lukeminen on?

Lukeminen on merkityksen abstrahoimista kirjoitetun tekstin perusteella: ääneen luettaessa on mukana lisäksi luetun ja ymmärretyn muuntaminen uudelleen puheeksi (ks. esim. Massaro 1978). Lukemiseen sisältyy useita mutkikkaita vaiheita, joista ensimmäinen on visuaalinen tiedon vastaanottaminen (silmän liikkeet, liikejaksojen koko ja hahmottuminen jne.; ks. esim. O'Regan 1979). Toinen vaihe, ortografisen koodin tulkinta, on teoreettisesti ongelmallisempi. Tutkijat ovat eri mieltä siitä, muunnetaanko ortografinen informaatio jonkinlaiseksi fonologiseksi tiedoksi, niin kuin fonologisen välittämisen teorian kannattajat väittävät, vai eteneekö ymmärtäminen suoraan ortografisen tiedon perusteella (ks. esim. Massaro 1978; Coltheart 1978; Levy 1978; McCusker & Hillinger 1981). Tapahtumaketjun viimeinen vaihe on viestin ymmärtäminen ja luku-tehtävässä sen muuntaminen jälleen puheeksi.

Viime vuosikymmenen aikana käsitys lukemisesta on muuttunut. Nyt korostetaan hierarkkisen tapahtumaketjun sijasta lukemistapahtuman kokonaisvaltaisuutta: on todettu, että sujuva lukija käyttää yhtäaikaaisesti monen eri tason vihjeitä ja keskittyy lukiessaan viestin merkityksen hahmottamiseen, ei niinkään yksityisiin sanoihin, kirjaimiin tai fraaseihin (ks. esim. Cooper & Petrosky 1976). Selvitettäväksi vielä kuitenkin jää, kuinka paljon sujuva lukija turvautuu tekstiin ja missä määrin taas käyttää hyväkseen olemassa olevaa sisäistä tietoa. Psykolinquistiset kokeet (ks. esim. Clark & Clark 1977) ovat osoittaneet, että kielellisen sanoman vastaanottaminen ei etene portaittain kieliopin eri tasojen kautta akustisesta tiedosta semanttiseen vaan hyödyntää usean tason tietoa samanaikaisesti.

Seuraavassa tarkastelemme sitä, miten tekstin tyyppi, rakenne ja vaikeusaste vaikuttavat lukusuoritukseen. Luennan vaikeudet heijastavat vastaanottamisen vaikeuksia, ja siten erilaiset virheet, takelut ja tauot kuvastavat eri sanojen tai tekstien vaikeutta. Jonkin verran käsittelemme myös lukijan asuinpaikan tai murretaustan vaikutusta lukusuoritukseen.

Tekstin vaikeuden käsite ei ole yksiselitteinen. Teksti voi olla käsitteellisesti vaikea: se voi käsitellä lukijalle vierasta aihepiiriä ja vaikeahahmotteisia kokonaisuuksia. Tällöin lukijan vaikeudet johtuvat ensyklopedisen tiedon puutteesta. Jos teksti on sanastollisesti vaikea, se voi sisältää runsaasti ani harvinaisia, vieraskielisiä tai hyvin pitkiä sanoja (vrt. Wiio 1975: 181—186). Vastaavasti helpon kielen sanastollisiin tunnuspiirteisiin kuuluisivat kotoperäiset, lyhyet ja tavalliset sanat. Grafofonologisesti vaikea teksti sisältää outoja kirjainyhdistelmiä ja suomen ortografisesta perusjärjestelmästä poikkeavia epäsäännöllisiä kirjoitusasuja. Teksti voidaan kokea vaikeaksi myös syntaktisesta syystä: rakenteet voivat olla kompleksisia, tekstissä voi olla paljon määritteitä tai teksti voi koostua pitkistä (yli 14—15 sanan) virkkeistä (tarkemmin ks. T. Itkonen 1972: 38—43; Platzack 1973; Wiio 1975: 181—186). On väitetty, että myös virkkeiden lyhyys saattaa haitata luettavuutta (Platzack 1973: 112—122; Sallinen-Kuparinen

1981: 180). Vaikealukuiselle tekstille on tyypillistä esimerkiksi heikko koheesio: tekstin ennustettavuus on vähäinen, ja lukusuoritukseen tulee väärien hypoteesien johdosta runsaasti epäsujuvia piirteitä. Seuraavassa keskitymme lähinnä niihin grafonologiisiin ja sanastollisiin vaikeuksiin, joita anglismikokeen tekstien luennassa ilmeni. Edellä esiteltyjen seikkojen perusteella voisi sekä poptekstin että sivistyssanatekstin odottaa osoittautuvan vaikeiksi, koska ne sisältävät runsaasti vierasperäistä sanastoa sekä vieraita kirjaimia, outoja kirjainyhdistelmiä ja vastaavasti vaikeita äännteitä ja äännejonoja.

3. Miten tekstin vaikeus heijastuu lukusuorituksessa?

Luenta voi olla joko sujuvaa tai epäsujuvaa. Siihen vaikuttavat useat yksilötekijät, esimerkiksi fysiologiset, sosiologiset, älykkyys- ja persoonallisuustekijät, sekä itse teksti (ks. esim. Vähäpassi 1977: 11—). Tässä esityksessä tarkastelemme vain tekstin ja joidenkin sosiolingvististen tekijöiden osuutta. Sujuvuus näyttää olevan ominaisuus, jota on subjektiivisesti helppo arvioida mutta objektiivisesti vaikea määritellä ja mitata (ongelmista ks. esim. Lehtonen 1981). Tekijöitä, jotka kuitenkin näyttävät selittävän luennan sujuvuuden tai epäsujuvuuden vaikutelmaa, ovat tauotus, puhenopeus, virheet ja lipsahdukset sekä erilaiset takeltelut. Takelteluja ovat toistot, äännähdykset, väärät aloitukset, änkytys ja korjaukset (Clark & Clark 1977: 263—264).

Hypoteesina oli, että vierasperäiset sanat aiheuttavat tekstiin epäsujuvuuksia, joiden määrä heijastaa lukijoiden lukuvaikeuksia. Tuloksista havaitaankin, että sivistyssanat, jotka eivät todennäköisesti kuuluneet lukijoiden aktiiviseen sanavarastoon, aiheuttivat enemmän epäsujuvuuksia kuin ns. nuorisolanglismit, joista ainakin osa kuului koehenkilöiden jokapäiväiseen kielenkäyttöön ja joiden käytöisyys (availability) oli siten suuri. Jos lukemiseen sisältyy olemassa olevan sisäisen tiedon hyväksikäyttöä, odottaisi sisäisen tiedon käyttöisyyden ja tekstin vaikeuden korreloivan keskenään negatiivisesti. Pelkän vierasperäisyyden ei siis pitäisi vaikuttaa lukusuoritukseen häiritsevästi. Jos nuorisolehtien tekstejä jäljittelevän poptekstin anglismit todella ovat nuorison omaa kieltä, ei sen luennassa odottaisi esiintyvän epäsujuvuuden tuntomerkkejä enempää kuin normaali-tekstiä luettaessa.

Seuraavassa tarkastellaan taukojen, virheiden ja takeltelujen määrää sekä puhenopeutta ja taukoprosenttia kunkin tekstin lukusuorituksissa. Taulukko 1 osoittaa taukojen, takeltelujen ja lukuvirheiden osuuden kussakin tekstissä. Taulukko 2 osoittaa kunkin tekstin luentasuoritusten lukuopeutta kokonaisnopeutena ja ääntönopeutena eli nopeutena, joka saadaan, kun taukojen osuus vähennetään luennan kokonaiskestosta. Lisäksi taulukosta 2 ilmenee taukojen keston prosentiosuus kunkin tekstin lukemiseen käytetystä kokonaisajasta.

Jaakko Lehtonen ja Hannele Heikkinen

TAULUKKO 1. Taukojen, takeltelujen ja lukuvirheiden useus eri teksteissä sataa sanaa kohti.

	Popteksti kpl	Normaaliteksti kpl	Sivistyssanateksti kpl
Tauot	23,0	18,0	24,0
Virheet	1,5	1,6	5,3
Takeltelut	2,2	2,3	7,1

TAULUKKO 2. Puhenoisuus, ääntönoisuus ja taukoprosentti eri teksteissä.

	Popteksti	Normaaliteksti	Sivistyssanateksti
Kokonaisnoisuus (sanaa minuutissa)	108	116	76
Ääntönoisuus (sanaa minuutissa)	144	140	99
Taukoprosentti	25	17	23

Tauot eivät puheessa ole yksinomaan epäröinnin tai lukemisen vaikeuksien osoittimia. Ylimalkaan tauot palvelevat kommunikaatiota sekä puhujan että kuulijan kannalta. Puhuja käyttää taukoja sekä puheen suunnitteluun että tekstin jaksotteluun. Taukoja puhuja tarvitsee myös sisäänhengitykseen; Goldman-Eislerin (1968: 96) mukaan siihen käytetään keskimäärin neljää taukoa viidestä. Hyvän lukijan tuntomerkki onkin muun muassa se, että hän on oppinut hengittämään siinä, missä teksti suo siihen mahdollisuuden: teksti määrää hengittämistä eikä hengittäminen tekstiä (ks. myös Miron & Brown 1971: 154). Tauot ovat tärkeitä myös kuulijan kannalta. On osoitettu, että tauot palvelevat kuulijaa antamalla hänelle aikaa viestin käsittelyyn (ks. esim. Butterworth 1980). Tauoilla, niiden määrällä, pituudella ja sijainnilla on siis kommunikatiivinen merkityksensä. (Tauotusta ja taukojen tehtäviä käsittelevästä laajasta kirjallisuudesta ks. Appel & Dechert & Raupach 1980.)

Taukojen määrä ja niiden pituus korreloivat sujuvuuden kanssa yleensä siten, että mitä enemmän ja mitä pitempiä taukoja luenta tai puhe sisältää, sitä epäsujuvammaksi se koetaan. Tekstin vaikeus lisää selvästi taukojen määrää (ks. esim. Pürschel 1975: 59). Puhenoisuus on kiinteästi sidoksissa tauotukseen. Puhe ei saa olla liian nopeaa eikä liian hidasta. Suomenkielisen normaalitekstin sujuva lukunoisuus näyttää olevan noin 120 sanaa minuutissa. Vapaan puhumisen nopeus on normaalisti hitaampi kuin lukunoisuus (Lehtonen 1978; 1979: 44).

Taulukosta 2 huomataan, että poptekstin ja sivistyssanatekstin taukoprosentit eroavat normaalitekstistä. Normaalitekstin taukoprosentti on 17, mikä vastaa täsmälleen aiemmissa tutkimuksissa sujuvasta tauotuksesta saatuja tuloksia (Lehtonen 1978). Aiemman tutkimuksen mukaan suomalaisten opiskelijoiden taukoprosentti englanninkielisen tekstin luennassa oli 31. Tämän tutkimuksen

Anglismit ja tekstin luettavuus

sivistyssanateksti ja popteksti sijoittuisivat taukoprosentin perusteella siis normaalin äidinkielen helpohkon tekstin ja vieraskielisen tekstin prosenttilukujen välille.

Koska ääntönopeus eli puhumisen »nettoaika» poptekstissä ja normaali-tekstissä on jotakuinkin sama mutta sivistyssanatekstissä huomattavasti lyhyempi, voimme päätellä, että lukijat ovat pitäneet sivistyssanatekstissä keskimäärin pitempiä taukoja. (Tässä tutkimuksessa ei yksityisten taukojen kestoja merkitty muistiin; kunkin tekstin nopeus ja taukoprosentti analysoitiin erityisellä tätä tehtävää varten suunnitellulla laskimella.)

Pitemmät tauot kuulija tulkitsee epäröinniksi usein silloinkin, kun ne sijoituvat lauseopillisesti hyväksyttävälle sananrajalle (ks. esim. Ruder & Jensen 1971: 120 ja Sallinen-Kuparinen 1981), ja niiden arvellaan olevan merkkejä puhujan kognitiivisesta toiminnasta (Kowal et al. 1975: 565—566). Nämä puheeseen epäsujuvuusvaikutelmaa aikaansaavat ja luennassa suhteellisen harvinaiset tauot osoittavat sivistyssanatekstin selvästi muita vaikeammaksi. Vierasperäiset sanat ovat lisänneet taukojen määrää myös poptekstissä, mutta sivistyssanatekstin pientaajuiset sanat ovat kasvattaneet myös taukojen keskipituutta.

Kun tarkastellaan virheiden ja takeltelujen määrää (taulukko 1), erottuvat popteksti ja normaaliteksti omaksi ryhmäkseen. Niissä on takelteluja ja virheitä selvästi vähemmän. Kaikki vierasperäiset sanat tekstissä eivät siis välttämättä lisääkään virheiden ja takeltelujen määrää. Takeltelujen ja lukuvirheiden jakauman perusteella ns. nuorisoanglismeilla ja sivistyssanoilla on puhujien leksikossa erilainen asema.

Taulukossa 3 on kukin koeteksti jaettuna riveille alkuperäisen koetekstin mukaisesti. Kullekin sananrajalle on merkitty siinä havaittujen taukojen määrä prosentteiksi muunnettuna. Kahtakymmentä prosenttia eli lukijoiden viidenestä suuremmat prosenttiluvut on kursivoitu. Yleensä puhuja jaksottaa tekstiä niin, että tauot sijoittuvat syntaktisten jaksojen välisille rajoille. Ns. epäröintitauot, jotka usein ovat myös kestoiltaan pitempiä, saattavat esiintyä myös kiinteästi toisiinsa liittyvien lauseenjäsenten välissä. On mahdollista, että suomalainen ei pidä välimerkkitaukoja yhtä tärkeinä kuin esimerkiksi englanninkielisen tekstin lukija. Tähän sopii myös Wiion (1975: 186) väite, että välimerkeillä ei ole suurta merkitystä suomenkielisen tekstin ymmärrettävyyden kannalta. Utriais (1981) mukaan suomenkieliset lukijat pitävät ruotsin- ja englanninkielisiä lukijoita vähemmän taukoja lauserajoilla. Virkkeiden rajalla suomenkielisen tekstin luennassa sen sijaan pidetään Sallinen-Kuparisen (1981: 177) mukaan tauko lähes aina. Suunnilleen samaa osoittavat myös tämän tutkimuksen tulokset: pisteen kohdalla tauotetaan hyvin usein mutta ei aina. Pilkku-tauotus vaihtelee huomattavasti enemmän: pilkun edellä piti tauon 17—85 % lukijoista virkkeestä riippuvasti. Välimerkkien mukaisen tauottamisen vaihtelu saattaa myös heijastaa sujuvien ja epäsujuvien lukijoiden eroja: esim. Sallinen-

Jaakko Lehtonen ja Hannele Heikkinen

Kuparisen (1981: 178) mukaan lukiolaiset jaksottelevat lukemaansa välimerkkien mukaan ammattikoululaisia tarkemmin. Tauot heijastavat myös kielen syntaktisia rakenteita ja niiden käsittelyä lukemisen aikana. Esimerkiksi attribuutin ja pääsanan välinen konstituenttirakenteen rikkova tauko rikkoo myös sujuvuutta. Tällaiset tauot ovat luennassa harvinaisia. Arajuuri (1980) on esitellyt myös tuloksia tekstin hahmottamisesta. Hänen mukaansa kirjoitettua tekstiä kopioitaessa kopioijat jaksottavat tekstiä samalla tavalla kuin lukijat. Kopioija sijoittaa jakson rajan usein juuri pisteen tai pilkun kohdalle tai syntaktisen jakson rajalle mutta harvoin esimerkiksi sanaliittojen jäsenten väliin.

TAULUKKO 3. Taukojen, taketelujen ja lukuvirheiden sijainti koeteksteissä. Taukojen prosentteiksi muunnettu määrä koko koehenkilöjoukon (150 henkeä) lukusuorituksissa on merkitty numeroin sananrajoille tekstin yläpuolelle. Lisäksi on kursivoitu numerot, joiden kohdalla yli viidennes lukijoista on pitänyt tauon. Rivin alapuolelle on sananrajalle merkitty lyhyellä viivalla sellaiset kohdat, joissa yli 10 % lukijoista teki lukuvirheen.

popteksti

Minusta tuntuu,¹ että¹ olen⁵ ihan⁰ tavallinen⁰ nuori.⁹⁰
 Yleensä⁷ päälläni³ on⁴⁷ college³-paita¹⁶ ja⁶ samettifarkut,⁸⁵
 jotka³ ostan⁴¹ Jeans⁵ and³⁹ Jackets¹⁵ -nimisestä¹ kaupasta⁸⁵
 lähellä²¹ City¹⁷ Centeriä.⁹² Kuuntelen¹³ popia⁹ ja³ luen¹⁶
 Helpiä¹² tai³ Introa.⁹⁶ Parasta² musaa² ovat² Abba,⁴¹ Baccara⁴⁶
 ja³² Bonnie¹⁴ Tyler.⁹³ Shokkibändejä⁷ kuuntelen⁵ diskossa.⁹³
 Kun² svengaa,³⁵ tulee⁶ fantastinen¹ fiilinki!⁹⁷ Toisaalta¹³
 kesällä⁵ on⁶ ihan¹⁵ o.k.⁴² grillata²² rannalla²³ selkäänsä³³
 ja³ kuunnella¹⁵ country-¹musaa.⁹⁷ Silloin¹ olisi⁷ mahtavaa⁹
 ohjata⁷ Jeepiä¹⁹ tai⁴⁵ Beach⁵¹-Buggya⁵⁴ pitkin¹² hietikoita.⁸⁹
 Joskus¹ illalla⁵ tulee⁹ mentyä¹⁶ publiin³ kaljalle,⁷¹ jos¹ ei¹³
 nimittäin⁹ missään ole⁷ bileitä.⁹⁷ Alaikäiset⁴ tietenkin⁷
 juovat¹ vain⁷ Spriteä⁵ ja⁵ kaakaota.⁸⁴ Schän on¹ selvä.

normaaliteksti

Liike-²⁶ säästö-¹⁴ ja¹ osuuspankit²⁸ yhdenmukaistavat¹⁴ tili-
 numeroidensa¹³ merkitsemisen.⁹⁶ Itse³³ tilinumerot³ eivät²
 muutu,⁵⁰ ainoastaan niiden⁷ merkitsemistapa.⁹⁷
 Asiakkaiden³³ esipainetut⁶ lomakkeet⁶ voidaan⁵ kui-
 tenkin¹⁰ sellaisenaan² käyttää¹ loppuun.⁹⁸
 Liike-²³ säästö-¹² ja osuuspankit⁵ ovat¹ sopineet,²⁵ että²¹
 tilinumeroon⁴ merkitään⁵ vain⁴ yksi¹³ väliviiva⁶⁶ ja³ näin¹¹
 ollen¹⁹ viimeisen⁷ numeron⁵ edessä¹⁵ ollut³ viiva⁵ jää pois.

sivistyssanateksti

– Alkoholintutkimuslaitoksen⁵⁹ – delegaatio²⁰ jätti⁴ hallitukselle⁷ muistion,⁵⁵ jossa³ kiinnitettiin² huomiota⁹ alkoholin²⁶ sekä¹¹ erilaisten⁴ huumeiden²⁷ kuten¹⁷ hashiksen¹⁵ ja¹⁶ – amfetamiinin³⁴ – lisääntyneeseen¹ käyttöön.⁹⁶ Myös³ muut³ lääkeaineet,¹⁷ kuten³³ kortisoni³¹ ja³⁸ – antihistamiini³⁵ olisi³⁰ delegation³ mukaan¹⁸ saatava⁹ tehostetun⁷ valvonnan¹ piiriin.⁹⁷ Eri¹ lääkeaineiden⁵ vaikutus²⁴ – ns.³³ – anabolisten¹⁴ – steroidien⁵¹ – eritykseen¹⁹ on¹¹ nykyään²¹ – analysoitavissa¹⁸ lähes⁵⁵ – simultaanisesti²⁷ laitoksen¹⁵ uuden⁶⁵ – alfanumeerisen²⁴ – mikroprosessorin⁶ avulla.⁹⁷ Lääkeainetta⁴ – nauttanut henkilö⁷ joutuu⁵ kokeessa¹⁰ ensin²⁵ Cooperin³ testiin,⁶¹ minkä² jälkeen¹⁴ mahdollisten²⁹ doping²-aineiden¹⁶ vaikutus¹⁶ mitataan¹¹ yksinkertaisesti²³ henkilön⁷¹ – betarytmin²¹ muutoksesta.⁹⁹ Delegaatio¹¹ ehdottaa muistiossaan¹¹ erityistä⁶³ – anti-alkoholismikampanjaa.⁷⁴ joka⁸ – toteutettaisiin¹³ – New¹ – Yorkiin²⁶ – sijoituneen⁶⁷ World⁴¹ Health⁶⁷ Centerin¹³ valvonnassa.

Myös yksityiset sanat selittävät taukojen sijoitusta sekä virheiden ja takeltelujen määrää. Sellaisissa lukukokeemme tekstien kohdissa, joissa sanan lukemisessa on runsaasti virheitä ja takelteluja, on myös taukoja, joko vaikean sanan edellä tai sen jäljessä. Tästä esimerkki on normaalitekstin sana *esipainetut*: 33 % lukijoista on pitänyt tauon sanan edellä, 19 % lukijoista on takelellut sanan lukemisessa ja 7 % on tehnyt virheen sanan kohdalla. Sivistyssanatekstin vaikeimpia sanoja taas oli *simultaanisesti*: 55 % lukijoista piti tauon sanan edellä, 36 % takelteli ja 26 % teki virheen. Virhekasautamaa ei kuitenkaan selitä yksi sana; myös sen edellä on vaikeuksia aiheuttavia rakenteita, esimerkiksi vokaalisoinnun vuoksi vaikeääntöinen *analysoitavissa*. Yleensä vaikeuksia aiheuttivat selvästi vieraskieliset sanat kuten *World Health Center* tai *Jeans and Jackets*. Kuitenkin sanan tuttuudella näyttää olevan merkitystä: esimerkiksi *New York* on ollut helpompi tuottaa kuin *Beach Buggy*. Sanan vieraskielisyys vaikeuttaa hahmottamista myös Arajuuren (1980) tulosten mukaan: ennen tai jälkeen vieraskielisen sanan pidetään kopiointitehtävässä mielellään pysähdys. Myös ääntämykseltään selvästi vieraskieliset sanat erottuvat omaksi ryhmäkseen (ks. Heikkinen 1981). Lukija ei pysty soveltamaan näihin sanoihin suomen kielen ortografisen koodin tulkintasääntöjä eikä myöskään suomen kielen ääntämyssääntöjä vaan joutuu turvautumaan oppimiinsa vieraan kielen sääntöihin. Tällaisen normaalista poikkeavan käsittelyn vaikeudet heijastuvat taukojen, takeltelujen ja virheiden määrässä ja sijainnissa.

Myös ns. sivistyssanat, esimerkiksi *alfanumeerisen*, aiheuttavat vaikeuksia. Jokaisen sivistyssanan edellä on poikkeuksellisen paljon taukoja, aina yli viidennek-

Jaakko Lehtonen ja Hannele Heikkinen

sellä lukijoista.¹ Näiden sanojen vaikeus ei niinkään näytä johtuvan yksin kirjoitusasun outoudesta. Esimerkiksi sanat *antihistamiini* ja *simultaanisesti* ovat taukojen, takeltelujen ja virheiden aiheuttajina kärkejoukkoa, vaikka ne sekä ääntämis- että kirjoitusasunsa puolesta ovat ongelmattomia. Myöskään sellaiset sanat kuin *svengaa*, *disko* tai *fiilinki* eivät ole aiheuttaneet poikkeuksellisen paljon epäsujuvuuksia eivätkä erotu normaaleista suomalaisista sisältösanoina omaksi ryhmäkseen. Sivistyssanan vaikeuden täytyy siis usein olla sen käsitteellisessä vaikeudessa tai muussa seikassa kuin yksin vieraassa kirjain- tai äänneaineksessa.

Sanojen pituus vaikuttaa osaltaan taukoihin ja takelteluihin (vrt. myös Sallinen-Kuparinen 1981: 227). Esimerkiksi sopii jo edellä mainittu *antihistamiini*: 38 % lukijoista on pitänyt tauon sen edellä, 51 % on takellellut sanaa lukiessaan ja 11 % on tehnyt lukuvirheen. Pitkien sanojen vaikeus ei välttämättä johdu niiden monikirjaimisuudesta; syynä saattaa olla myös sananmuotojen morfosyntaktinen kompleksisuus (tästä ks. Karlsson 1979). Joissakin tutkimuksissa on todettu (ks. esim. Vasić 1976: 614), että sanan morfologisen rakenteen harvinaisuus lisää lukuvirheitä. Tässä tutkimuksessa aiheuttivat paljon virheitä sellaiset sanat kuin *lisääntyneeseen*, *sijoittuneen* tai *esipainetut*.

Lukuvaikeuksia aiheuttavat myös lyhenteet, joista sivistyssanatekstissä oli esimerkkinä lyhenne *ns*. Tämän lyhenteen edellä piti 24 % lukijoista tauon, 27 % takelteli ja peräti 33 % teki virheen. Sallinen-Kuparinen (1981: 228) on myös todennut lyhenteiden vaikeuden ja tulkitsee tämän johtuvan suomen kielen pleonastisesta taivutusjärjestelmästä: epäsujuva lukija ei hahmota tarpeeksi pitkiä kokonaisuuksia osatakseen valita lyhenteelle oikean taivutusmuodon. Lyhenteet poikkeavat myös normaalista ortografisesta koodista, ja kirjoitusasun monitulkintaisuus saattaa aiheuttaa vaikeuksia suomen yksiselitteiseen kirjoitukseen tottuneille lukijoille.

Myös kielioppisanat näyttävät aiheuttavan tiettyntyyppisiä vaikeuksia. Niihin nimittäin keskittyy huomattavan paljon lukuvirheitä. Esimerkiksi poptekstin virheitä aiheuttanut sana oli *tai*: tekstin ensimmäisen *tai*-sanan luki 22 % *ja*-sanaksi, ja toisen kohdalla 8 % lukijoista teki saman virheen. Tällaiset virheet osoittavat, että lukija ei kiinnitä paljon huomiota tekstin funktiosanoihin vaan tuottaa ne oman sisäisen synteesisä perusteella. Cole, Jakimik & Cooper (1978) ovat puheen varjostuskokeissa saaneet samankaltaisia tuloksia: tekstin varjostaja havaitsee huonosti funktiosanoihin tarkoituksellisesti sijoitettuja virheitä.

Virheet, takeltelut ja tauot eivät kuvasta tekstin käsittelyn vaikeuksia samalla tavoin. Lukuvirheetkin voi jakaa sujuviin ja epäsujuviin. Sujuviin kuuluvat

¹ Poikkeuksena oli sana *alkoholi*. Sana näkyy omaksutun suomen kielen puhujien normaaliin sanavarastoon, eikä sitä enää koeta vieraaksi. Ääntämyksessä lähes kaikilla oli muoto /alkoholi/ pro /alkkohooli/.

sellaiset luvuvirheet, jotka eivät aiheuta taukoja tai takelteluja sanan ympäristöön, vaikka aiheuttavatkin poikkeaman tekstiin. Tällaiset virheet keskittyvät funktiosanoihin, eivätkä ne heijastele tekstin tai sanan vaikeutta. Ne kuvastavat ehkä paremmin lukemisen luonnetta: silloin tällöin lukija sisäistä hypoteesia rakentaessaan tekeekin vääriä ennusteita. Jos ennustettu ja todellinen sana ovat yhtä hyväksyttäviä, ne vaihtuvat helposti keskenään. Näin on käynyt kun *tai* on luettu *ja*-sanaksi ja sana *lähellä* sanaksi *läheltä*.

Epäsujuvia ovat virheet, jotka aiheuttavat ympäristöönsä taukoja ja takelteluja. Tällaisia luvuvirheitä aiheuttavat sivistyssanat ja kirjoitusasun poikkeuksellisuudet mutta myös pitkät ja rakenteeltaan kompleksiset sanat. Epäsujuviksi virheiksi on ehkä myös luokiteltava spontaanille puheellekin ominaiset ennakointi- ja perseveraatiivirheet: ääntämys /gollege/ /kollege/:n sijasta, /muos muut/, jossa /myös/-sanana /y/ on muuttunut /u/:ksi seuraavan sanan takavokaalisuuden vaikutuksesta, /bändejä kuundelen/ ilmauksesta /bändejä kuuntelen/ tai /on ihan ookoo grillata grannalla/ pro /on ihan ookoo grillata rannalla/. Tämänäyttävät virheet selittyvät peräkkäisten rakenteiden samankaltaisuudesta: samankaltaisia mutta kuitenkin hiukan erilaisia rakenteita on vaikea hallita lähekkäin, ja ne samanlaistuvat helposti (yhteyksistä virhesanontoihin ks. Heikkinen 1979).

4. Murteen ja asuinpaikan vaikutus lukusuoritukseen

Tutkimuksen testauspaikkakunnat pyrittiin valitsemaan niin, että ne edustaisivat erilaisia maantieteellisiä murrealueita ja toisaalta heijastaisivat pääkaupunkiseudun vaikutuksen määrää. Lahti on eteläsuomalainen kaupunki, joka toisaalta on alttiina pääkaupunkilaisvaikutteille mutta toisaalta sijaitsee eri murrealueella kuin Helsinki. Kuopio taas sijaitsee savolaismurrealueen sydämessä ja etäällä pääkaupunkiseudusta. Hypoteesina oli, että asuinpaikalla voi olla merkitystä anglismeihin suhtautumisessa, niiden tuttuudessa ja niiden käytössä.

Joitakin kiintoisia eroja eri seuduilta kotoisin olevien lukijoiden välillä tulikin ilmi. Esimerkiksi lahtelaiset lukijat (hämäläiseltä murrealueelta kotoisin olevat) käyttivät kaikissa teksteissä lukemiseen eniten aikaa! Myös ääntönopeus eli puhumisen »nettonopeus» ilman taukojen osuutta oli hämäläisillä johdonmukaisesti alhaisin. Erot ovat paikoin myös tilastollisesti merkitseviä. Näyttää siis siltä, että vanha käsitys hämäläismurteiden verkkaisuudesta ei ole kokonaan vailla pohjaa. Taukojenkin määrän suhteen hämäläiset erottuvat omaksi ryhmäkseen: he pitivät eniten taukoja joka tekstissä, ja ero savolais- ja helsinkiläismurteiden lukijoihin oli normaalitekstissä myös tilastollisesti merkitsevä. Myös luvuvirheitä ja takelteluja oli hämäläisillä yleensä jonkin verran enemmän. Sivistyssanatekstissä hämäläisillä on selvästi enemmän takelteluja ja luku-

Jaakko Lehtonen ja Hannele Heikkinen

virheitä: kun hämäläislukijan teksti sisältää keskimäärin 8,0 takeltelua, ovat luvut helsinkiläisillä 5,6 ja savolaisilla 6,3; erot ovat tilastollisesti melkein merkitseviä.

Mitä testin tulokset sitten kertovat kielellisten muutosten ja innovaatioiden leviämisestä? Tutkimuksen yleisraportissa todettiin (ks. Sajavaaran ja Lehtosen artikkelia tässä vihkossa), että helsinkiläiset ymmärsivät anglismeja selvästi savolaisia paremmin. Toisaalta helsinkiläiset suhtautuivat anglismeihin savolaisia negatiivisemmin. Savolaiset näyttävät pyrkivän muuta helpommin myös mukauttamaan uudet lainasanat ja vierassanat omaan taivutus- ja ääntämysjärjestelmäänsä. Tästä ovat merkinä myös eräät äänteelliset ilmiöt, muun muassa järjestelmällisempi /p/:n käyttö /b/:n sijalla kuin muissa murreryhmissä (ks. Heikkinen 1981). Murteen käyttö ja vieraiden äänteiden korvaaminen kotimurteen äänteillä kai myös tunnetaan paljon hyväksyttävämmäksi kuopiolaisten opiskelijoiden piirissä. Savolaisten suhteellinen sujuvuus ja virheiden ja takeltelujen vähyys saattavat siis heijastaa oikeastaan kielen äänne- ja muotojärjestelmän käytön varmuutta ja kykyä sopeuttaa oudotkin sanat omaan kielenkäyttöön.

Kokeen »normaalitekstin» kaltainen teksti, joka sisältää vain kielen ortografian perusjärjestelmään kuuluvaa ainesta, on todellisuudessa harvinainen. Uudempiin lainakerroksiin kuuluvia sanoja esiintyy enemmän tai vähemmän kaikissa niissä teksteissä, joiden kanssa kielenkäyttäjä joutuu päivittäin tekemisiin. Sanat, jotka ovat lukijalle tuttuja ja jotka esiintyvät niiden tulkintaa tukevassa kontekstissa, eivät sinänsä ole ongelmana yksin vieraan kirjoitusasun ja outojen kirjainyhdistelmien takia. Koska kieleen tulviva uusi aines voi vain osittain mukautua suomen ortografiseen ja fonologiseen perusjärjestelmään, uudet kielenkäyttäjät omaksuvat myös kirjoitukseen ja ääntämiseen uusia normeja. Ei ole myöskään näköpiirissä mitään, mikä merkitsisi tällaisen vieraan aineksen vähenemistä suomenkielisessä tekstissä. Mutta se kielenaines, joka mm. nuorisolehtien antaman kuvan mukaan olisi tämän päivän nuoren omaa kieltä, osoittautui tekstin luettavuuden perusteella ainakin tämän tutkimuksen koehenkilöille normaalia yleiskielistä asiategstiä vaikeammaksi ja vieraammaksi.

LÄHTEET

- APPEL, G., DECHERT, H. ja M. RAUPACH 1980: A selected bibliography on temporal variables in speech. Tübingen, Narr.
- ARAJUURI, Y. 1980: Lukutapahtuma, tulkintajakso ja tekstin luettavuus. Lisensiaatintyö. Tampereen yliopisto.
- BUTTERWORTH, B. 1980: Evidence from pauses in speech. — B. Butterworth (toim.), *Language production*, Vol. 1. Speech and talk. New York, Academic Press.
- CLARK, H. H. ja E. V. CLARK 1977: *Psychology and language*. New York, Harcourt Brace Jovanovich.
- COLE, R. A., JAKIMIK, J. ja W. E. COOPER 1978: Perceptibility of phonetic features in fluent speech. — *Journal of the Acoustic Society of America* 64: 1 s. 44—56.
- COLTHEARTH, M. 1978: Lexical access in simple reading tasks. — G. Underwood (toim.), *Strategies of information processing*. London, Academic Press.
- GOLDMAN-EISLER, F. 1968: *Psycholinguistics. Experiments in spontaneous speech*. New York, Academic Press.
- HEIKKINEN, H. 1980: Kielellisistä virhesuorituksista. — K. Suomi (toim.), *Fonetiikan päivät* — Oulu 1979. Oulun yliopiston fonetiikan laitoksen julkaisuja 1. Oulu.
- 1981: Vierassanat ja nykysuomi. Esitelmä 10. Fonetiikan päivillä Tampereella 20.—21. 3. 1981.
- ITKONEN, T. 1972: Kielenhuoltomme tehtävistä. — *Virittäjä* 76 s. 278—297, 371—387.
- KARLSSON, F. 1979: Automatic morphological segmentation of Finnish word forms. — *Papers from the Conference on General Linguistics*, Seili 29.—30. 8. 1979. Suomen kielitieteellisen yhdistyksen julkaisuja 3. Turku.
- KOWAL, S., O'CONNELL, D. C., O'BRIEN, E. A. ja E. T. BRYANT 1975: Temporal aspects of reading aloud and speaking: three experiments. — *American Journal of Psychology* 88 s. 549—569.
- LEHTONEN, J. 1978: On the problems of measuring fluency. — M. Leiwo ja A. Räsänen (toim.), *AFinLA:n vuosikirja 1978*. Jyväskylä.
- 1979: Speech rate and pauses in the English of Finns, Swedish-speaking Finns and Swedes. — R. Palmberg (toim.), *Perception and production of English: Papers on interlanguage*. AFTIL 6. Publications of the Department of English. Åbo Akademi. Turku.
- 1981: Problems of measuring fluency and normal rate of speech. — J.-G. Savard ja L. Laforge (toim.), *Proceedings of the 5th Congress of AILA*, Montreal, August 1978, s. 322—332.
- (toim.) 1978: *Puheentutkimuksen menetelmät*. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 17. Jyväskylä.
- LEVY, B. A. 1978: Speech analysis during sentence processing: reading and listening. — *Visible Language* 12: 1 s. 3—26.
- MASSARO, D. 1978: A stage model for reading and listening. — *Visible Language* 12: 1 s. 3—26.
- MCCUSKER, L. X. ja M. L. HILLINGER 1981: Phonological recording and reading. — *Psychological Bulletin* 89: 2 s. 217—245.
- MIRON, M. S. ja E. BROWN 1971: The comprehension of rate incremented aural coding. — E. Foulke (toim.), *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*, University of Louisville, s. 133—161.
- O'REGAN, K. O. 1979: Saccade size control in reading: Evidence for the linguistic control hypothesis. — *Perception and Psychophysics* 25: 6 s. 501—509.
- PLATZACK, C. 1973: *Språket och läsbarheten*. Lund, Gleerup.
- PÜRSCHHEL, H. 1975: *Pause und Kadenz: Interferenzerscheinungen bei der englischen Intonation deutscher Sprecher*. Tübingen, Narr.
- RUDER, K. F. ja P. J. JENSEN 1971: Speech pause duration as a function of syntactic junctures. — E. Foulke (toim.), *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*, University of Louisville, s. 119—133.
- SALLINEN-KUPARINEN, A. 1981: Koululaisten kerronnan ja luennan parametrejä. Lisensiaatintyö. Jyväskylän yliopisto.

Jaakko Lehtonen ja Hannele Heikkinen

- UTRIAINEN, I. 1981: Placing of unfilled and filled pauses in native and non-native English and in native Finnish and Swedish. Pro gradu -työ. Jyväskylän yliopisto.
- VASIC, S. 1976: The association between reading aloud and various forms of speech behaviour. — G. Nickel (toim.), Proceedings of the 4th International Congress of Applied Linguistics Vol. 1 s. 411—423.
- WIIO, O. A. 1975: Ymmärretäänkö sanomasi? Helsinki, Weilin & Göös.
- VÄHÄPASSI, A. 1977: Lukutaidon rakenteesta ja vaihtelusta peruskoulun kolmannella luokalla lukuvuonna 1973—1974. Kasvatustieteiden tutkimuslaitoksen julkaisuja 283/1977. Jyväskylän yliopisto.

Anglicisms and text readability

Jaakko Lehtonen and Hannele Heikkinen
(Jyväskylä)

The Jyväskylä Anglicism project included a reading test in which 150 students from different areas of Finland read three texts. The short texts varied in style: the »popular text» imitated the style of popular magazines and included numerous words of foreign origin but colloquial in style (eg. *pub*, *fantastinen* 'fantastic' etc.). The »normal text» was a notice from a bank to its clients and did not include foreign words. The »foreign word text» imitated the style of a newsletter and employed several words of foreign origin and formal register (eg. *delegaatio* 'delegation'). The reading performances were recorded on tape and analyzed both aurally and instrumentally. This report includes results obtained from the analysis of pauses, speech rate, reading errors and other disfluencies.

The aim of the study was to find out to what extent textual difficulty, especially lexical and graphophonological difficulty, is reflected in the reading process. It was the hypothesis of the authors that those words, whether foreign or native, which belong to the active vocabulary of the readers do not cause difficulties while words that are unfamiliar and infrequent do. The results seem to support this hypothesis. The »foreign word text» was obviously the most difficult one: it was read at a much slower rate than the other two and the number of reading errors and disfluencies was radically higher. The number and average length of pauses was also clearly higher. At the lexical level, the items that seemed to be the most difficult

(i.e. that had more pauses immediately before and after them, or were read more disfluently or erroneously) included words that were clearly (e.g. orthographically) foreign (e.g. *Jeans and Jackets*). Also fairly new and infrequent loanwords (e.g. *simultaanisesti* 'simultaneously') were disfluent. Long words, both native and foreign, caused difficulties as well (e.g. *lisääntyneeseen* 'to the increased ...' or *antihistamiini* 'antihistamine'). Abbreviations were also difficult, probably because they violate conventional orthographic decoding rules and have to be learned as separate instances. Function words caused a lot of reading errors but not disfluent reading. These fluency errors show that the reader does not rely only on the input information but constructs hypotheses and predictions, based on his internal knowledge, while he reads.

The test also yielded some sociolinguistically interesting results. The speakers came from three towns: Helsinki, the capital; Lahti, a southern industrial town near Helsinki; and Kuopio, an eastern town far from Helsinki. The speakers from Kuopio showed least foreign influence and adapted the foreign words more readily into their native and dialectal idiom, both morpho-syntactically and phonologically. The students from Helsinki, as predicted, used non-native words or sounds more fluently and consistently. Interestingly, the speakers who came from Lahti were most disfluent of all. This is probably due to the fact that they are influenced by the prestige features of the

Helsinki dialect and feel compelled to use them, but obviously lack the means to do so. On the whole, the test clearly showed that the foreign loanwords do have an effect on various levels of Finnish. However, the foreign origin as such does not cause difficulties in the production of a word: the primary factors seemed to be

unfamiliarity, low frequency and formality. The popular anglicisms, which were part of the students' everyday vocabulary, did not seem to cause difficulties and also seemed to have a greater tendency to adapt to the morphological and phonological rules of Finnish.