Lund University, Dept. of Linguistics
Working Papers 50 (2002), 63 – 78

Some lexical features of immersion pupils' oral and written narration

Ulla Laurén
Department of Nordic Languages, University of Vaasa, Finland

1 Introduction

1.1 *Background*

In a current research project, I am studying how immersion pupils structure oral and written narratives in Swedish. It is a longitudinal study of the second-language development of pupils in the first two pioneering classes of a Swedish immersion programme in Vaasa, Finland. In this programme, the pupils' first language is Finnish while their immersion language is Swedish.

The pupils in this study started their immersion into Swedish in the autumn term of 1987 and 1988 respectively. They were then six years old and their first year of immersion took place in an immersion kindergarten a year before they began attending an immersion school. The programme during these first years formed an experimental project with a continuous follow-up throughout the pupils' lower and higher grades of the comprehensive school (for a survey of the research results, see Laurén Ch. 1999).

The immersion programme was started in 1987 as the first of its kind in Finland. The Canadian immersion programme has served as a model, but the Swedish immersion programme in Finland has acquired its own specific character. It is based on certain fundamental conditions and prerequisites (for more details, see Laurén Ch. 1999), and the following principles are central to it:

– The immersion language is a non-dominant language in the region. This is the case in Vaasa, where in 1996 73% of the inhabitants spoke Finnish, 26% spoke Swedish, and the remaining 1% another language (Herberts & Laurén 1998:41).

– Immersion is started early, at pre-school age, and as total immersion, i.e. only Swedish is used for instruction for the first few years.

– The proportion of the pupils' native language increases continuously so that after some five years, the immersion language (Swedish) and the native language (Finnish) occupy an equal part of the teaching. Learning the second

language is not allowed to take place at the expense of the native language, but bilingualism is meant to be an enrichment for the individual.
– The immersion programme is voluntary and open to all. It is not, however, meant for pupils who have some knowledge of the second language prior to beginning school. The child should not have, e.g., a Swedish-speaking parent or nanny.
– Immersion is directed by a conscious pedagogy, which emphasises the active role of the pupil in the process of learning. The language is the means, not the object of learning.
– The immersion programme aims at multilingualism. A third and sometimes a fourth language will already come into the picture at the lower level of the elementary school.

1.2 *The material, aim and method of the project*

In my project, I study the immersion pupils' progress in the development of their second language, Swedish, in comparison with corresponding age groups whose first language is Swedish. With regard to part of the material, the written narratives, I also compare the immersion pupils with Finnish-speaking pupils in Finnish schools where Swedish is taught in the traditional way. It should be possible to ascribe the differences between these two types of pupils not only to temporal and individual factors but also to didactic differences. The traditional teaching of Swedish implies that Swedish is taught *as a subject* a limited number of periods a week. In immersion, teaching takes place *in* Swedish, allocating a definite number of hours to each language, and applying definite didactic principles based on the pupils' independent work and responsibility for their own learning.

The material based on oral narration was collected in grade 5. Both immersion groups told a story derived from pictures in Mayer 1969 (the so-called Frog Story test). Later, in grade 8 and grade 9 respectively, the same two groups wrote a story, also based on pictures, dealing with two boys' adventures on a beach. The material for the orally related story in grade 5 was also given to a control group of Swedish-speaking pupils in the corresponding grade of a Finland-Swedish school. The material for the written story in grades 8 and 9 was given to two control groups from each of the corresponding grades. One group consisted of Swedish-speaking pupils in a Finland-Swedish school, the other group consisted of Finnish-speaking pupils with Swedish as their so-called A-language in a school for Finnish-speaking pupils. These pupils received traditional teaching of Swedish as their second national language in grades 3–6, two periods a week per school year. In

grades 7–9, Swedish was taught as the A-language 8 periods a week in all during those three years, usually 3 periods a week in grades 7 and 9 and 2 periods a week in grade 8.

Figure 1 shows the distribution of the material on the basis of oral and written narration respectively and according to the grade of the immersion groups and the control groups. The acronyms (SB1, SB2 etc.) included in the figure will below be used indicating the different groups concerned.

Oral picture-based narration in Swedish, grade 5 (two different age groups)
Immersion groups
– grade 5/SB1 (1993); 19 pupils (9 girls + 10 boys)
– grade 5/SB2 (1994); 21 pupils (9 girls + 12 boys)
Swedish control groups
– grade 5/SV1 (1993); 19 pupils (12 girls + 7 boys)
– grade 5/SV2 (1994); 18 pupils (8 girls + 10 boys)

Written picture-based narration in Swedish, grades 8 and 9
Immersion groups
– grade 8/USB2 (1997); 17 pupils (8 girls + 9 boys)
– grade 9/USB1 (1997); 19 pupils (9 girls + 10 boys)
Swedish control groups
– grade 8/USV8 (1997); 17 pupils (8 girls + 9 boys)
– grade 9/USV9 (1997); 19 pupils (10 girls + 9 boys)

Finnish control groups (Swedish as A-language)
– grade 8/UFI8 (1997); 17 pupils (8 girls + 9 boys)
– grade 9/UFI9 (1997); 19 pupils (9 girls + 10 boys)

Figure 1. Distribution of the material with regard to immersion groups and control groups.

I have analysed the oral and the written narratives on different levels. On a general story grammar level I have studied the plot and episode structure, that is, how well the immersion pupils build up the oral and the written narratives with a number of components needed to create coherence of content in the narratives. The model I have used is based on Labov 1972, Rumelhart 1975, Lindberg, Juvonen & Viberg 1990:171–172, Berman & Slobin 1994,

Trabasso & Rodkin 1994, Bamberg & Marchman 1994 and Knapp 1997:63–92. My model is presented in greater detail in Laurén 2001.

At the syntactic level, I have analysed, for instance, the use of connectors that introduce clauses or sentences. The use of certain connectors of this kind is typical of oral narration, and I have checked whether the connective expressions typical of oral narration can also be found in the written narratives.

I will complete the picture of the immersion pupils' second-language usage by a lexical analysis. In this article, I will present my results regarding one part of the lexical analysis, namely the part concerned with lexical density and lexical variation. At this stage of the investigation, statistical calculations of significance have not yet been performed. At the end of the article, I will by the aid of a qualitative analysis of the written narratives discuss how useful the global measures of lexical density and lexical variation are for the comparison of various types of language learning and stages of language acquisition.

2 Lexical density and lexical variation

2.1 *Measuring lexical density and lexical variation*

Vocabulary is essential for language acquisition and language use. Its extent and quality with regard to variation and the use of content words and form words are part of the linguistic competence, which in different studies has proved to be covariant, for instance, with grammatical competence (Grönholm 1994:308, Cummins et al. 1990:122–123). To measure the vocabulary in the two types of narrative I have used two different global measures, lexical density and lexical variation.

Lexical density refers to the proportion of lexical words of all the words of a text, and the measure is also considered applicable to texts of varying lengths (Linnarud 1986:47). Lexical density has been used in different studies of the composition of vocabulary from the perspective of both first and second language. I have used the measure in the same way as Linnarud 1986 and Korkman 1995, and give the quotient of lexical words per 100 running words. Lexical words are meaningful words, i.e. nouns, main verbs, adjectives and adverbs, which end in -t in Swedish. Auxiliary verbs thus remain outside the category of lexical words. By marking off auxiliaries from main verbs, I have made it possible to compare lexical density with other investigations. Both with regard to the oral and the written material, I have observed the same principle of division as is used, for example in Korkman 1995:155, with reference to a definition of classification as a list in Teleman 1974:48, 231. As a result, the following Swedish verbs are counted as auxiliaries: *vara, bli, ha,*

vilja, kunna, måste, komma...att, skola, göra and *få*. This means that these verbs, according to Teleman 1974:231, are classified as auxiliaries even when used in the function of main verbs. This categorisation differs from the traditional definition of auxiliary verbs, for instance, according to Thorell 1973.

Lexical variation indicates the ratio between the number of distinct words and all words. The ordinary type/token value is sensitive to variations in the length of the text, and consequently different measures have been proposed where the length of the text is not equally decisive. For instance, OVIX (Hultman & Westman 1977) and other formulas of lexical variation are based on logarithmic numbers or square root values (for additional information, see Pitkänen & Kohonen 1984:173–178). In my analysis I have used the OVIX formula $V = N^{(2-N^k)}$ (Hultman & Westman 1977:264; Hultman 1993:62), where V stands for the number of running words, N for the number of different words (so-called lexical, material or full words), and k is a constant calculated on the basis of logarithmic values for V and N. As for lexical words I have counted one and the same graphic word. So Swedish *var,* for example, is one and the same graphic word irrespective of its part of speech and meaning, whereas *var* and *vara* are two different graphic words. Different inflected forms of one and the same word thus constitute different lexical words. In Hultman and Westman's study of secondary-school pupils' essays, the OVIX value proved to serve well as a measure of lexical variation and correlated positively with the marks given for the essays. Irrespective of the length of the text, weak essays received low OVIX values and good ones high OVIX values. The OVIX values varied in the secondary-school pupils' essays between less than 60 for small lexical variation and more than 70 for great variation (Hultman 1993:62). The interpretation of the OVIX values must, however, be seen against the background of text-external factors (e.g. first-language and second-language vocabulary respectively, social group or sex) as well as text-internal factors (e.g. text genre and style).

2.2 *Lexical density and lexical variation in the oral narratives*

From the oral narratives I have, when counting the number of running words in the text and considering lexical density and variation, left out items indicating hesitation (*hmm, ää, öhh, öhmm* and others) as well as slips of the tongue and repetitions (e.g. SV206P: o *glasskå* [//] glasflaskan går sönder; *<sen gick de ett>* [//] oo *<sen po>* [//] sen letar pojken i den stora eken).

Lexical density can also be used as a measure of colloquial vs. literary language. Halliday 1989:64 points out that the more literary a text is, the higher is the proportion of lexical words to running words. Ure 1971 found in

a study of oral and written texts, that the oral texts showed a lexical density of below 40% while the density of written texts usually exceeded 40%. The most important distinction in Ure's study was, however, whether it was a question of interaction or not. The occurrence of verbal as well as non-verbal response from an interlocutor tends to lower the degree of lexical density. Hyltenstam 1988:73 has, in a study of bilingual Finnish–Swedish and Spanish–Swedish as against monolingual Swedish secondary-school pupils' oral and written production, found similar differences as Ure between spoken and written language. The lexical density of oral texts varied between 35.9 for the monolingual Swedish control group and 38.3 for the bilingual Finnish–Swedish group of pupils, while the lexical density of the written language of these groups was about 45 per cent.

The average values of lexical density (Table 1) are both for the immersion groups and for the control groups around 35%. In other words, they are in agreement with Ure's and Hyltenstam's investigations of spoken language. The lexical density of the immersion pupils is, nevertheless, somewhat higher than the values represented by the control groups. This can possibly be interpreted as a lower degree of "nativeness" in the immersion pupils. The pupils in the control group are more likely to pad out their narration with various kinds of form words like pronouns, adverbs and prepositions than to use meaningful lexical words. Individual variations are, however, considerable in both the immersion groups and the control groups.

Table 1. Lexical density in the oral narratives of the immersion pupils in grade 5 (SB1 and SB2) and the Swedish-speaking control pupils (SV1 and SV2).

Group	Lexical density	Range of distribution
SB1	37.4	28.8 – 50.6
SV1	36.9	31.1 – 51.3
SB2	37.6	28.6 – 52.4
SV2	33.0	25.9 – 45.3
SB tot. (grade 5)	37.5	28.6 – 52.4
SV tot. (grade 5)	34.5	25.9 – 51.3

Table 2 shows that the Swedish-speaking control groups display a clearly wider lexical variation than the immersion groups. The two immersion groups

are very similar. They show on the average the same lexical variation, and the range of distribution is also the same in the two groups. The OVIX values of the oral narratives are low compared with the OVIX values of written essays both in this study and in comparison with other studies of written material (see section 2.3).

Table 2. Lexical variation (OVIX) in the oral narratives of the immersion pupils in grade 5 and the Swedish-speaking control pupils (SV1 and SV2).

Group	OVIX	Range of distribution
SB1	35.3	29.1 – 45.1
SV1	42.1	34.7 – 49.0
SB2	35.3	29.8 – 45.2
SV2	39.4	33.7 – 54.2
SB tot. (grade 5)	35.3	29.1 – 45.2
SV tot. (grade 5)	40.8	33.7 –54.2

2.3 *Lexical density and lexical variation in the written narratives*

The lexical density and the lexical variation of the written narratives have been measured according to the same method of calculation as the oral narratives. In the essays, linguistic items wrongly written as two words or as one word have been corrected since they directly affect the number of words. On the other hand, misspellings and other errors have not been corrected. The lexical density, i.e. the proportion of meaningful words of all words, will appear from Table 3. Table 3 shows that lexical density is clearly greater in the written than in the oral material from grade 5. This is in entire conformity with previous results regarding the difference between oral and written material (Ure 1971, Hyltenstam 1988:73, Halliday 1989:64). Except for the Finnish control group in grade 8, the lexical density in my written research material lies between 40 and 50 per cent, which generally agrees with the lexical density of about 45 per cent represented by native and near-native Swedish secondary-school pupils' written language (Hyltenstam 1988:73).

As for the material produced by the different groups I have studied, certain differences appear in Table 3. The oral material for the Swedish-speaking control groups shows lower lexical density than that of the immersion pupils (Table 1, section 2.2), and this tendency is further emphasised in the written material. In both grade 8 and grade 9, the Swedish-speaking control groups display the lowest lexical density. With regard to the narratives produced by

pupils with Swedish as their native language, it is obviously natural for the meaningful lexical items to be surrounded by various non-meaningful form words providing cohesive textual links. If the Swedish-speaking control groups are seen as a norm, both the immersion groups and the Finnish-speaking control groups with their clearly higher lexical density deviate from this pattern. In grade 9 the immersion group and the Finnish-speaking control group do not markedly differ from each other. In grade 8, on the other hand, the immersion pupils display much lower values regarding lexical density than the Finnish-speaking control group, which shows very high average lexical density with some extremely high values within the group. The Finnish-speaking control group produced very short essays, which contributes to the increase of lexical density, since new referents are introduced more frequently than in a longer narrative.

Table 3. Lexical density in the written narratives in grades 8 and 9 of the immersion pupils (USB1, USB2), the Swedish-speaking control pupils (USV8, USV9), and the Finnish-speaking control pupils (UFI8, UFI9).

Group and grade	Lexical density	Range of distribution
Grade 8		
USB2	46.7	39.1 – 54.4
USV8	42.0	36.6 – 46.8
UFI8	62.5	41.1 – 86.1
Grade 9		
USB1	48.5	39.5 – 56.9
USV9	40.8	35.5 – 50.7
UVI9	47.6	41.9 – 55.6

Lexical variation has been calculated in the same way as for the oral material, i.e. according to the OVIX formula. In calculating OVIX for the essays, linguistic items wrongly written as two words or as one word have been corrected since the wrong form would directly affect the number of words and thus the OVIX value. Compared with the oral narratives (see Table 2, section 2.2) in grade 5, the written narratives in grades 8 and 9 of both the immersion pupils and the Swedish-speaking control pupils show markedly greater lexical variation (Table 4). This increase can be partly attributed to the higher age of the pupils and to the difference between oral and written production.

Table 4. Lexical variation (OVIX) in the written narratives in grades 8 and 9 of the immersion pupils (USB1, USB2), the Swedish-speaking control pupils (USV8, USV9), and the Finnish-speaking control pupils (UFI8, UFI9).

Group and grade	OVIX	Range of distribution
Grade 8		
USB2	53.0	43.5 – 71.9
USV8	56.1	47.0 – 61.8
UFI8	43.9	30.9 – 68.7
Grade 9		
USB1	49.2	38.1 – 58.9
USV9	53.1	39.8 – 63.1
UFI9	50.9	37.1 – 91.7

Table 4 further shows that the Swedish-speaking control groups in both grade 8 and grade 9 display the widest lexical variation. The immersion group in grade 8 shows a clearly higher average OVIX value than the Finnish-speaking control group. In grade 9, on the other hand, the immersion group shows a somewhat lower average value than the Finnish-speaking control group. One pupil in the Finnish-speaking control group has, however, an extremely high OVIX value of 91.7, which affects the average value.

2.4 *Qualitative analysis of lexical density and lexical variation in the written narratives*

Lexical density (LD) and OVIX represent purely mathematical values, and it may therefore be advisable to examine essays with extremely high values qualitatively, and to compare essays with similar values between the different groups. Examples (1) – (4) are essays with the lowest LD values and varying OVIX in their respective groups. Of the four essays, essay (2), written by a Swedish-speaking control pupil, has the lowest lexical density, and essay (4), written by a Finnish-speaking control pupil, the lowest lexical variation.

The lowest LD value in the two immersion groups and a relatively low OVIX value are exemplified in the following essay (grade 8):

(1) *En farlig överraskning* (heading)
Det var en vacker sommardag. Hans och Erik var i stranden. I himmelen fanns det inga moln. Det var bara solen som sken. Det var en jättevarm dag. Efter en stund blev pojkarna hungrika, och dom började äta. Erik och Hans var båda ihop med en flicka, och ville överraska dom, men visste inte hur. Dom började fundera, och efter en stund såg Erik att uppe i bergen växte det vackra blommor. Dom visste att det skulle vara

farligt att klätra upp till bergen men dom bestämde att göra det för flickornas skull. Dom gick längre och längre bort från blommorna, för att titta om det skulle finnas vackrare blommor där i lite längre bort. Till sist kom pojkarna till en så smal plats att om dom skulle ännu röra sig, skulle dom falla.

Pojkarna började ropa hjälp av andra människor, och till sist fick dom hjälp. En helikopter kom med ropen till pojkarna och dom blev räddt. (USB202F; LD 39.1; OVIX 45.2)

Example 2 is an essay written by a pupil from the Swedish-speaking control group in grade 8. It has the lowest LD value in the group. OVIX is below the average value of the group.

(2) *Olyckan på stranden* (heading)
Det var en solig junidag. Jag och min kompis var nere på stranden och solade oss. Vi hade med en picknikkorg som vi smaskade i oss. När vi hade ätit upp all maten gick vi runt på stranden och letade efter fynd. Vi hittade några vackra snäckor, men det riktiga fyndet var några väldigt vackra blommor som mycket sällsynta.

Vi bestämde oss för att plocka hem några åt mamma och pappa. Vi klättrade upp en bit men sen slapp vi varken högre upp eller ner. Nu får vi problem tänkte båda två och började ropa på hjälp. Det tog ett bra tag förran någon hörde oss men till slut kom det en familj som var ute och gick på stranden, de undrade vad vi gjorde där och hur vi hade kommit opp dit, vi berättade för de och sade att vi inte slipper ned igen. Pappan i familjen ringde efter sjöbevakarna som var där med en helikopter innom några minuter. En man skickades ner ur helikoptern där vi stod band fast oss och så skickades vi opp i helikoptern. Vi tackade sjöbevakarna och pappan, sen for vi hem.
(USV804P; LD 36.6; OVIX 54.3)

Essay (3) by a Finnish-speaking control pupil in grade 9 has the lowest lexical density in the group. OVIX is above the average value of the group.

(3) *Pojkarna på stranden* (heading)
Peter och Stefan är bästa vänner. En dag gick de till en strand att ta solen och så vidare, när det var så vackert vädret. Solen skiner och det var +25° varmt. Pojkarna hade litet mat med och, när de blev hungriga, de åt smörgåsar och apelsiner och drack limonaden. När de hade ätit, gick de och promenera lite på stranden.

De såg en mycket hög plats framför dem, och där växte vackra blommor. När de gick dit uppe, de sliper inte bort därifrån. Peter skrikade: -hjälp!!! De satt där uppe tio minuter och sen nån människor rindge till polisen. De kom nästan genast med helikopter och hämtade pojkar bort därifrån. Pojkarna var glada när de slipper tillbaka hem.
(UFI902F; LD 41.9; OVIX 53.9)

In the Finnish-speaking control group in grade 8, essay (4) has the lowest lexical density and at the same time the lowest OVIX in the group.

(4) *Två pojkar en strand* (heading)
Två pojkarna hade på en strand de hade föd korg och de äta smörgås och dricka saft och efter några tim de ämna gå till berg och de samla blomma och den de ta när och den en män ringa en polis och helicopter till gå rädda de och den de hade skydd. (UFI810P; LD 41.1; OVIX 30.9)

What narratives (1) – (4) have in common is thus low lexical density in comparison with the other narratives in their respective groups. Examples (1) and (2), written by an immersion pupil and a Swedish-speaking control pupil respectively, both show OVIX values below the average of their groups. The two narratives are written in a colloquial everyday style. Example (3), written by a Finnish-speaking control pupil, has the lowest lexical density of the Finnish control group, which yet corresponds to the average of the Swedish control groups and has an OVIX close to (2). The Finnish-speaking control pupil in (3) also uses a comparatively fluent and varied language in comparison with the Finnish control pupil in (4). In narrative (4), the lexical density corresponds to the average value of the Swedish control groups, while OVIX is much lower than in (1) – (3). Judging from these four texts, the LD value does not describe the linguistic quality too well, especially if one wishes to compare groups with different linguistic proficiency with one another; OVIX seems to serve better in this respect.

Examples (5) – (8) are all essays with high lexical density values within their respective groups. Regarding all immersion pupils, essay (5) has the greatest lexical density, and its OVIX is above the average values for immersion pupils in both grade 8 and grade 9.

(5) *Hans-Åke och Börje ligger på stranden.*
De badar i solen. Sen äter de smörgåsar och dricker First Choice Cola. Sen ser Hans-Åke vackra blommor på klippan. Pojkarna klättrar på klippan. De samlar blommor. Sen märker de att de kan inte klättra tillbaka. Börje ropar Hjälp! Palle är på stranden med gossarna från messi. Han ringer till kustbevaktningen. Snart kommer Håkan med helikoptern och räddar pojkarna. (USB112P; LD 56.9; OVIX 55.8)

Example (6) is an essay that represents the highest value of lexical density in the Swedish control groups. Its OVIX value represents the second highest individual lexical variation in the Swedish control groups, in grade 8 as well as grade 9.

(6) *Spluukus Blommus* (heading)
En varm sommardag när solen sken bestämde sig Göran och Jörgen att de skulle fara till Tallholmen för att söka efter den sällsynta blomman Spluukus Blommus som bara växte i detta område. Blomman kunde de sedan sälja till ryssland för mycket pengar.
När de kom till ön började de genast söka efter den extremt sällsynta och fridlysta blomman. Efter tre timmars intensivt sökande, tog de en paus för att äta reissu-mies-smörgåsar och dricka mummon sekamehu-saft. Plötsligt såg Jörgen en mystisk blomma högt uppe på ett berg, det var Spluukus Blommus. De packade fort ihop deras matsäck och började klättra upp på berget.
När de äntligen nådde till blomman började de plocka den, pojkarna var så insatta i plockandet att de inte märkte folkmassan på stranden. Plötsligt såg Göran den arga folkmassan, oo-nej det var blommornas vänner som ville slå ner dem i en mörk gränd för att de plockade den sista Spluukus blommus.
Pojkarna försökte fly, men förgäves. En gammal moffa hade ringt polisen och de såg en helikopter komma mot dem.
Tio minuter senare satt Göran och Jorgen ledsna i Helikoptern.
Göran och Jörgen fick fyra års fängelse för brott mot naturskyddslagen. Göran begick självmord efter det andra året.
(USV914P: LD 50.7; OVIX 60.8)

In the Finnish control groups, example (7) represents an extreme value as far as lexical density is concerned. The same pupil's OVIX value is among the lowest in the two Finnish control groups.

(7) *Pojkar på stanen* (heading)
Juha och Mika solar sig på stanen, hunglig poikar laga mat, Juha titta blommor, poikarna blogga vagra blommor. Poikarna vara red, man ser poikar. Man ringa polishelikopter. Polishelikopter regna man, polishelikopter gar poikar. (UFI817P; LD 86.1; OVIX 37.0)

The Finnish control group in grade 9 shows no such extreme values such as the corresponding Finnish control group in grade 8. Example (8) displays the first group's highest lexical density with an OVIX of 58.3, which at the same time constitutes the second highest OVIX value of the group.

(8) *Tom och Daniel har varit på stranden hela dag. Dagen var varmt och de njöt av sit livet. Toms mamma hade lagit goda smörgåsar och pojkarna åt dem. Daniel såg blommor uppe i backen. Pojkarna ville samla dem och de steg upp till backen.*
Menniskorna på stranden såg pojkarna och en man ringde till policen och de låvade att hjälpa pojkar.
Tom och Daniel var glada när helikopter kom. Allt slutade bra.
(UFI914F; LD 55.6; OVIX 58.3)

What texts (5–8) have in common is, therefore, a high LD value in comparison to other narratives in the same group. The immersion pupil's narrative in (5) is relatively simple, but new referents are often introduced, which increases the number of content words, and the lexical variation is relatively wide. In the Swedish control pupil's narrative (6) there is a rich variety of various attributive words, which increases both lexical density and lexical variation.

The extreme values of the Finnish control group in grade 8 are obviously due to the fact that some of the pupils are still at a low level of linguistic development, as example (7) with a lexical density of over 80 per cent confirms. In this example, pronouns are almost wholly lacking (*man* is in the context a noun and not a pronoun), and so are other text-linking words. Some content words are repeated (*poikar, polishelikopter, Juha*), which increases the number of meaningful words, but the text is atomistic and difficult to understand. OVIX is low because of these repetitions and because of the lack of connectors.

Example (8) has about the same degree of LD as (5) but it surpasses that of the Swedish-speaking control pupil's essay (6). Compared with (6), text (8) is still very simple although the texts have an almost equal OVIX. An explanation of the differences in linguistic quality is the difference in length. Lexical density and lexical variation do not, after all, give the same result for texts of different length.

According to Hultman and Westman 1977:56–60, an index of word variation is a good criterion for the linguistic quality of an essay. The relation between an index of lexical variation and standard is the most significant relation they have found between a single language variable and standard. Even if the measure is created in order to make a comparison between texts of various lengths possible, it is a weakness that it seems to function differently for different stages of linguistic competence. Pupils speaking their native language cannot, therefore, out of hand be compared with learners at different stages of their linguistic development.

3. Final comments

The results of the analysis which I have previously carried out on the content structure of plot and incidents (Laurén 2001; Laurén in progress), show that immersion pupils in the fifth grade are quite capable of orally telling a story in Swedish based on a previously unfamiliar sequence of pictures. With regard to aspects that have been studied, immersion pupils in the eighth and ninth grades show further progress when it comes to writing a story in their second

language similarly based on an unfamiliar sequence of pictures. The written narratives show that the immersion pupils develop in the direction of a more literary use of clause-initial connectors in main clauses. They still, however, seem inclined to overuse certain connectors, probably as a result of influence from their first language.

As far as lexical density and variation described in this article are concerned, it can be concluded that immersion pupils follow the same lines of development as the Swedish-speaking pupils in the control groups in progressing from oral narrative in grade 5 to written narrative in grades 8 or 9. The number of content words increases and the pupils vary their choice of words more in the written than in the oral narratives. This development can be reasonably explained in terms of the difference between spoken and written language, but also of natural linguistic and cognitive development during the school years. Moreover, on the part of the immersion pupils, this development can be explained as a result of an increased vocabulary in the second language.

At the same time, it is obvious that the measuring of lexical density and lexical variation must be done in a context in order to be adequately interpreted. It is normal for an oral narrative to show low lexical density and lexical variation. With regard to written narratives, those who use the language as their mother tongue and those who use it as their second language cannot be compared offhand. The degree of second-language proficiency attained and the individual variation, especially among those who have not yet advanced very far in their second-language development, seem to play an essential part with regard to lexical density and variation.

References

Bamberg, M. & V. Marchman. 1994. 'Foreshadowing and wrapping up in narrative'. In R.A. Berman & D.I. Slobin, *Relating events in narrative: a crosslinguistic developmental study*, 555–590. Hillsdale N.J.: Lawrence Erlbaum Associates, Publ.

Berman, R. A. & D. I. Slobin. 1994. *Relating events in narrative: a crosslinguistic developmental study.* Hillsdale N.J.: Lawrence Erlbaum Associates, Publ.

Cummins, J., B. Harley, M. Swain & P. Allen. 1990. 'Social and individual factors in the development of bilingual proficiency'. In B. Harley, P. Allen, J. Cummins & M. Swain (eds.), *The development of second language proficiency*, 119–133. Cambridge: Cambridge University Press.

Grönholm, M. 1994. 'Sanojen ja muotojen oppiminen – tasatahtiako toisessa kielessä'. In L. Laurinen & M.R. Luukka (toim.), *Puhekulttuurit ja kielten oppiminen*, 307–319. AFinLAn vuosikirja 1994. Suomen soveltavan kielitieteen yhdistyksen (AFinLA) julkaisuja 52.

Halliday, M.A.K. 1989. *Spoken and written language*. Oxford: Oxford University Press.

Herberts, K. & Ch. Laurén. 1998. *Urbes linguas suas habent*. Vasa: Skrifter utgivna av Svensk-Österbottniska samfundet 59.

Hultman Tor G. 1993. 'Hur gick det med OVIX?' I: *Språkbruk, grammatik och språkförändring. En festskrift till Ulf Teleman 13.1.1993*, 55–64. Lund: Lunds universitet.

Hultman, T.G. & M. Westman. 1977. *Gymnasistsvenska*. Lund: Liber Läromedel.

Hyltenstam, K. 1988. 'Lexical characteristics of near-native second-language learners of Swedish'. *Journal of Multilingual and Multicultural Development 9*, 1&2, 67–84.

Knapp, W. 1997. *Schriftliches Erzählen in der Zweitsprache*. Tübingen: Max Niemeyer Verlag.

Korkman, C. 1995. *Tvåspråkighet och skriftlig framställning. En undersökning av tvåspråkiga elevers uppsatser i den finlandssvenska grundskolan*. Helsingfors: Studier i nordisk filologi 74. Skrifter utgivna av Svenska Litteratursällskapet i Finland 593.

Labov, W. 1972. *Language in the inner city. Studies in the black English vernacular.* Philadelphia: University of Pennsylvania Press.

Laurén, Ch. 1999. *Språkbad. Forskning och praktik*. Vaasa: Vaasan yliopiston julkaisuja. Tutkimuksia 226. Språkvetenskap 36.

Laurén, U. 2001. 'From oral to written story. On the development of immersion pupil's structuring of a story'. In S. Björklund (ed.), *Language as a Tool. Immersion Research and practices*, 295–314. Proceedings of the University of Vaasa. Reports 83.

Laurén, U. 1994. *Tvåspråkiga och enspråkiga skolelevers skriftliga produktion. Performans och kreativitet*. Vasa: Acta Wasaensia 37. Språkvetenskap 4.

Laurén, U. In progress. *Andraspråkstillägnande genom språkbad. Studier i berättelsestruktur samt lexikal och grammatisk kompetens hos språkbadselever*. University of Vaasa.

Lindberg, I., P. Juvonen & Å. Viberg. 1990. 'Att berätta på två språk – en studie av tvåspråkiga barns återberättelser på första- och andraspråket'. In U. Nettelbladt & G. Håkansson (utg.), *Samtal och språkundervisning. Studier till Lennart Gustavssons minne*, 169-192. Universitetet i Linköping: Linköping Studies in Arts and Science 60.

Linnarud, M. 1986. *Lexis in composition*. Lund: CWK Gleerups.

Mayer, M. 1969. *Frog, Where are you?* New York: Pied Paper Printing.

Pitkänen A.J. & V. Kohonen. 1984. *Johdatus kvantitatiiviseen kielentutkimukseen ja alan ATK-sovelluksiin*. Helsinki: Gaudeamus.

Rumelhart, D.E. 1975. 'Notes on a schema for stories'. In D.G. Bobrow & A. Collins, *Representation and understanding*. New York: Academic Press, Inc.

Teleman, U. 1974. *Manual för grammatisk beskrivning av talad och skriven svenska*. Lund: Lundastudier i nordisk språkvetenskap. C:6.

Thorell, O. 1973. *Svensk grammatik*. Stockholm etc.: Esselte Studium.

Trabasso, T. & Rodkin, P.C. 1994. Knowledge of goal/plans: A conceptual basis for narrating *Frog, where are you?* In R. A. Berman and D. I. Slobin, *Relating events in narrative: a crosslinguistic developmental study*, 85–106. Hillsdale N.J.: Lawrence Erlbaum Associates, Publ.

Ure. J. 1971. 'Lexical density and register differentiation'. In G.E. Perren & I.L.M. Trim (eds.), *Applications of linguistics*, 443–452. Cambridge: Cambridge University Press.

Viberg, Å. 1992. 'Språkutveckling i förskolan och skolan. Svenska som första- och andraspråk från 6 till 16 års ålder'. In S. Hellberg m.fl. (utg.), *Svenskans beskrivning 19*, 289–298. Lund: Lund University Press.

Viberg, Å. 1993. 'Introduktion: Projektet Tvåspråkighet i skolan'. In Å. Viberg, V. Tuomela & P. Bergman, *Tvåspråkighet i skolan. Bas-rapport 1. Rapporter om tvåspråkighet*. Stockholm: Stockholms universitet. Centrum för tvåspråkighetsforskning.

ulla.lauren@uwasa.fi

Lund University, Dept. of Linguistics
Working Papers 50 (2002), 79 – 84

Naming Amongst 6 Year Old Children

Anna Simonsen

The Rehabilitation Unit of Folkhälsan, Helsinki, Finland

1 Introduction

1.1 Purpose

The purpose of this paper is to present the first steps and the first results of a long term study, the CoPS project, about early assessment of dyslexia in children using the translated computer program CoPS (Cognitive Profiling System) and other screening tests. CoPS has been developed as an early identification method established by reference to cognitive precursors, such as phonological awareness and memory (Singleton et al 2000). The final aim in the long-term study is to analyse the language among monolingual Swedish speaking and Swedish/Finnish bilingual Finnish children and the specific symptoms language impairment has in these groups. Furthermore, it will be studied how these characteristics develop, and which characteristics are significant in developing dyslexia. The children have now been screened at the age of 5 years (N 400); some of these have been thoroughly assessed at the age of 6 years (N 130). The children's language skills will be reassessed at the end of the second school year including assessment of reading and writing skills.

Phonological processing plays an important role in literacy development as well as language abilities involving vocabulary and even grammar (Catts et al. 2000). Nauclér and Magnusson 2000 have done a longitudinal study of language impaired children vs. control children, evaluating spoken as well as written language capacity. They found a significant difference between the groups in a semantic word retrieval task, whereas no significant difference could be seen in a phonological retrieval task. Also, the scores in naming and language comprehension differed significantly between the groups. In written language, there was no difference in decoding scores, but there was a difference in reading comprehension between the groups. These results suggest that there are difficulties in different semantic processes in language impaired children which are influent in speech as well as written language.