

Speeding up vocabulary acquisition by immigrant children

Anne Vermeer, Tilburg University, the Netherlands (www.AnneVermeer.com)
Iceland, Friday, May 12th 2006

0. Introduction

The lack of vocabulary in the second language is the biggest problem for immigrant children. This talk deals about vocabulary learning and teaching: first some theoretical notions: what do we mean by 'word', why are words so important, how are words learned; and secondly: aspects of teaching: selection of words, learning activities and exercises. Next, I will report on a research project on speeding up the vocabulary of immigrant children, and finally, show you some examples of teaching materials.

The lack of vocabulary in the second language is the biggest problem for immigrant children. I showed in my first talk a graph of the vocabulary sizes in elementary school of low SES Dutch native children, and immigrant children.

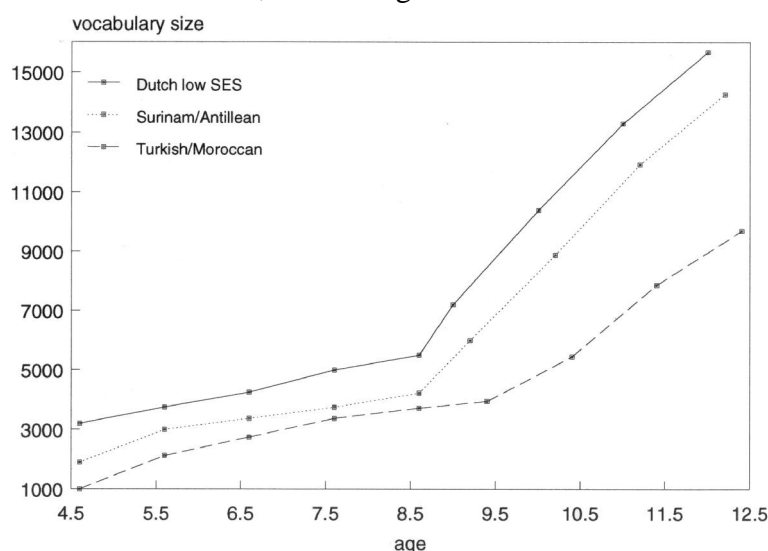


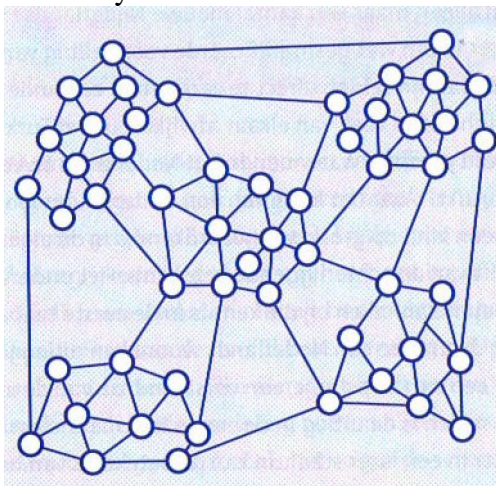
Figure 1. Absolute size of receptive vocabulary of three ethnic groups in primary education

Figure 1 gives an indication of the size of the receptive vocabulary of three groups: the solid line indicates Dutch lower Social Economic Status children. The dotted line in the middle represents Surinamese and Antillean children, children with colonial backgrounds who speak mostly Dutch and sometimes another language at home. The lower dashed line indicates the Turkish and Moroccan migrant workers' children, who mostly speak the mother tongue at home. The horizontal axis shows the age of the children, from 4 to 12 years old. On the vertical axis the absolute number of words known is displayed. The data are based on studies Ludo Verhoeven and I did with two proficiency tests, administered to more than 4,000 children: the first one in age groups (4- to 9-year-olds in grades 1-5); the second in grades 5 to 8 (9- to 12-year-olds). As can be seen in Figure 1, the differences between Dutch lower SES children, on the one hand, and Turkish and Moroccan children, on the other, are dramatic. I will give three examples. (1) At the age of 6, when children learn to read and write, Turkish and Moroccan children have a vocabulary size that is much smaller than that of their native Dutch classmates entering Kindergarten, two years before. (2) At the age of 9, when they are confronted with science texts, Turkish and Moroccan children know fewer Dutch words than five-year-old Dutch children - approximately 3,500 words. (3) At the end of primary school, at the age of 12, these migrant children's vocabulary in Dutch is only two thirds of their Dutch classmates' vocabulary - roughly 10,000 versus 15,000 words.

Over the last twenty years, vocabulary has been upgraded as a component of language proficiency, both in L1 and L2 language acquisition. In the past, there has been an overvaluation of morphological and syntactical skills, even in languages that have only little morphology at all, such as English. That overvaluation had also to do with the traditional structuralist view on language of main stream linguistics (the Generative school of Noam Chomsky). Various morphological procedures like conjugation (I go, and I went) or declension, or compounding (as in *baðherbergi* (n.) - bathroom or *svefnherbergi* (n.) – bedroom) were traditionally thought to be derived of the root word. These words were supposed to be formed by language speakers on the basis of a rule. It was assumed that forms like *geese*, or *went*, or *bathroom* would be understood by language learners on the basis of their knowledge of grammar and root words. Nowadays, the view on morphological rules in language acquisition and learning has been changed dramatically. I come back to this point later. Anyway, the knowledge of words is now considered the most important factor in language acquisition and school success – in part due to its close ties with text comprehension. Words are the carriers of meaning. Without knowledge of words, understanding sentences or texts is not possible. It is exactly as Lewis wrote: ‘Language consists of grammaticalised lexis, not lexicalised grammar’.

1. What is ‘word knowledge’ and how are words learned?

Knowledge of words is multidimensional and incorporates various types of knowledge. Words are sometimes represented as nodes in a network. These nodes are interconnected with other nodes or words. For instance, *table*, *chair*, *to sit*, *to eat*, *plate*, *spoon*, *fork* and *knife*, etcetera are together in the domain ‘dining room’ (*borðstofa*) or ‘dinner’. That is a dense network, words that come up together easily: if you hear or say the one word, the other words are ready to jump out of your mouth. Especially *fork* and *knife* have special bands. That is because you hear these two words often together in one sentence.



The denser a network around a word, the greater the knowledge of that word, and the easier to retrieve words from the same theme or domain.

Next, we have to distinguish between concept and label. The label is the superficial form, the spoken or written form of a concept, what we usually call ‘word’. The concept is what lies underneath. The concept includes all associations, ideas, feelings and visual aspects that are connected to that word. Here again you can imagine all those associations as a kind of network around that concept. The more associations, feelings, ideas and visual and other aspects, are linked up with that word, the better you know that word. Such a network is build up by encountering the word in various contexts. Take the concept SHEEP. You know how a sheep looks alike (visual, white or black; these black sheep are mine). You know their sound

(beeh or meeh, depending on the language they speak), and other features (wool, milk) and experiences (you've been to a farm once with your class, for instance). And for you, the label 'sheep' is very well connected to all these aspects and experiences. But Arabic and Turkish immigrant children have experiences like *Ait el Kebir*, the feast of sheep at the end of the Muslim fast (the month that they don't eat nor drink during daytime). All of these children KNOW of course the concept SHEEP. However, they do not know the LABEL 'sheep', but they know the label 'kebir' in Arabic or 'koyun' in Turkish. In learning Icelandic, such a label is not really acquired when you encounter only once the word 'kind'. The network has to be built again, by coming across many times the label 'kind' in various linguistic contexts. I looked it up, searched the internet for a typical Icelandic sheep, I had to laugh because the label 'kind' means 'child' in Dutch, I wrote it down, wrote it down again in a slide in powerpoint. So now I know that the concept 'sheep' has the label 'kind' in Icelandic. It is important to realize that you have to encounter a word in many forms, in many different contexts, many times, to finally remember and retrieve the word in its different forms, so you can use it yourself.

So, if an Turkish speaking child has heard 'kind' only once or twice, there is only a very thin line between the concept and the word. He needs a denser network around the word 'kind' to remember it! And repetition is even more necessary, as he doesn't hear the difference between 'sheep' and 'ship' very well, between /i/ and /ie/. ('There is a ship in my barn')

I think it is necessary to repeat here again that a word has different word forms, especially in a morphological complex language as Icelandic. For instance, forms of *kind* like *kindar* and *kindur*. All these word forms have a node in the network. They are all stored in the mental lexicon in the meaningful, verbal context in which they are used. So, the child has to encounter all these different forms in the appropriate context in order to acquire them. They learn these forms together with the environment in which they are used. So, for me, I learned quite easy that *bokarinnar* was a genitive, because I understood that 'the end of the book was the most exciting'. (*Lok bókarinnar eru mest spennandi*). And of course, it can be helpful and it is useful to learn the different morphological rules behind these forms, but first, the learner has to encounter them and use them in co-occurrence with the grammatical patterns they fit in, as in *Lok bókarinnar eru mest spennandi*.. Only after that, learning to analyse and exercise with the different word forms in the different contexts can promote language learning (so called *focus on forms*). But the teaching materials for second language learners has to be quite different from the materials used for first language learners, because the former need a lot more repetition of good examples, and they do not have native intuition. And because, for example, gender is so a prominent feature in Icelandic, you cannot build exercises on native intuitions about gender. Second language speakers do not have that intuition, they do not have gender differences, or their gender system is not the same (for instance, *book* as well as *sheep* are feminine, in Icelandic, however, in Dutch they are neuter).

The question how well a word is known, is to a great extent dependent on the amount of connections between these different aspects, dependent on the density of the network. The denser the network, the better the word is known, and the easier the word can be retrieved from your mental lexicon in your head. Moreover, the easier you can guess the meaning of other usages of the word, for instance the use of 'sheepish'. 'Richard was looking *sheepish*'.

Finally, knowing a word is always a question of 'more or less' knowing. Word knowledge is a continuum from receptive to productive knowledge, that gradually emerges. First, you don't even recognize the word, then you recognizes the word: ('I've seen it before'). Then you recognizes the context of use, but not the meaning ('It has something to do with'). Then, you recognises the meaning of the word: first global knowledge in an

association, next, only one meaning in only one context, and thirdly, more meanings in more contexts. More productive is already that you know the meaning but you cannot retrieve the word (known as 'the-tip-of-the-tongue'). And finally, you use the word in the context, in only one context or more contexts, and you are able to describe the meaning of the word: visual and functional features, synonyms or antonyms, or a verbal definition. Then, you have full productive knowledge of the word.

Receptive

- n doesn't recognize the word
- n recognizes the word ('I've seen it before')
- n recognizes context of use, but not meaning ('has something to do with ...')
- n recognizes meaning of the word:
 - first global knowledge in an association,
 - next, only one meaning in only one context,
 - third more meanings in more contexts.
- n knows meaning but cannot retrieve the word ('the-tip-of-the-tongue').
- n uses word in context
 - in only one context
 - or in more contexts
- n is able to describe the meaning of the word:
 - visual and functional features,
 - synonyms or antonyms
 - gives verbal definition.

productive

2. Teaching vocabulary: which words, how many words, and how to teach

2.1 Which words?

As I showed earlier, if a language learner encounters one unknown word in a sentence of ten known words, he can guess the meaning of that unknown word. But, if two or more words out of ten are unknown, the language learner will be at a loss. So, you need to know a lot of words. But which words do you choose for teaching?

Some words in a language are much more frequent than others. In the question which words have to be learned first, the frequency of words in a language plays an important role. The first 3000 to 5000 words in a language cover already about 80 per cent of a text. These words are a kind of basic vocabulary, to be taught first. Because, once a learner knows these, say, 4000 words, he can guess the meaning of other words more easily, and learn these other words more easily. Once children know these 4000 words, you saw in the vocabulary figure, the lines show a significant increase.

It is for curriculum developers quite important to know: which are these words? I will not go into detail here, but in designing curriculum content in vocabulary for Dutch as a second language we made use of two kinds of lists. The best and most valid way to find out the most important words is to gather at least 2 million words in spoken and written input to children. That costs a lot of blood, sweat, and tears, and money, too. I can tell you, because I did that research some ten years ago. We administered oral input of teachers to children, texts of picture books, reading materials, science texts, history, arithmetic, etcetera, etcetera, and counted all word forms. In this way, we have now a frequency list of about 25.000 words. Teachers and curriculum developers use this list to find out the difficulty of a text, and to see what words are too difficult for a particular learner, etcetera. They can use the device freely, it is available on my homepage *Woordwerken* (Working with words).

A much more simpler way is the following. Take a list of about 6000 words. I did that once in developing a frequency list for Moluccan Malay. All words in that list were evaluated by about 80 teachers. They had to mark each word in the list by YES or NO on the question whether, 'I think this word has to be known by a 6 year old child', or 'no, I don't think that it has to be known'. In this way, the words of which ALL teachers say that is has to be known,

are the most important. And the list ends with the words of which only a few teachers say that these have to be known. This is a simple and quite reliable way to know which are the 3000 to 4000 most important words. I think it is worthwhile to develop such a list for children learning Icelandic as a second language, too.

2.2 How many words?

As I said, it is for immigrant children very important to learn these 4000 words as quick as possible. But how many words can be learned in a certain amount of time?

According to the literature, 6 to 8 words can be taught and learned in a lesson of 20 minutes, offering explicitly the words in different contexts and repeating regularly. If you do that 4 times a week, 1000 to 1500 words in a year must be possible. We tried to find out this in the research I will tell you about in the third part of my talk. Before that, there are some vocabulary teaching principles that I want to tell you about first, some principles we incorporated in our research project, too.

2.3 How to teach?

In the vocabulary teaching literature it is mentioned that a word needs at least seven times, to be exposed for a learner before it is acquired: four times the first time in different contexts and situations, and then repeated later on, three times at least. A kind of formula is used for that: $4+1+1+1$. For instance, you read a story to the children about a farm; the sheep passes by four times in various sentences in that story; and after some days you repeat that story again, a week after again a second time, and the third repetition is for instance in another story.

A second often mentioned aspect of vocabulary teaching is the way in which the child attach the appropriate meaning to a word. According to the literature, four steps in teaching have to be followed:

- * Preparation
- * Semantization
- * Consolidation
- * Evaluation

The first step, preparation, means that the teacher offers a context, a domain, in showing a picture, asking questions like 'where do you think the story in this book is about?' and thus, make all kind of associations in the mental lexicon of the child possible around the context or domain where the story is about. The purpose is to awake the network in the child's mental lexicon around the concepts and words he already knows. In this way, the new words can be attached more easily into the network of the child.

In the second step, semantization, the appropriate meaning of the new word is attached to the word, for example, by showing a picture, telling what it does or works or is used for, how it looks alike, showing the action, and so on.

In the third step, consolidation, the specific word is firmly attached in the memory, by using it in other contexts, by asking what it does or works or is used for, or how it looks alike, or the child can show what it is, or the children play the story while the teaching is reading aloud, and so on.

The fourth step is evaluation, in which, after some time, the teacher checks whether the child has really acquired the specific word.

I hope the relation is clear between what I've told about *the knowledge* of a word (the mental lexicon with all kinds of nodes and interrelated networks between those nodes) on the one hand, and on the other hand, the *learning and teaching* of a word (building up for the child a

very dense network in order for him to acquire the word). Thus, in the context of a story, or a theme, and in using visual elements like photo's, pictures, visual objects, in using other senses like taste, feeling, and by 'show and tell'-lessons with the children, the teacher tries to build up as many traces in the network as possible.

3. Speeding up vocabulary

3.1 Design of the research project

On the basis of the preceding considerations, my colleague Rene Appel from the University of Amsterdam, and I, designed an experimental program to speed up the acquisition of Dutch vocabulary by migrant children in the first four grades of primary education, Starting in Kindergarten 1 till the end of grade 4. In the Netherlands Kindergarten is part of primary school, which has eight grades, the first two in Kindergarten. It was hoped that at the end of the four-year period, the Dutch vocabulary of the migrant children would be equivalent to that of their native peers, with approximately 5,000 to 6,000 words (receptive knowledge). In order to reach this goal, each year at least 1,200 new words, the 'target words', were presented to the children, in vocabulary teaching materials, about 4. 600 different words in all. Teachers gave 30 minutes of instruction four times a week with materials prepared for children learning Dutch as a second language, sometimes within the classroom, and sometimes outside the classroom in small groups. One of our goals was also to 'revitalize' courses Dutch as a second language outside the classroom.

In general, these curriculum materials for young children have a format in which a certain story has a central position. About half of the new words each year were part of these regular curriculum materials for Dutch as a second language. The other half of the new words in each year was presented in children's picture books. The story, which is supported by pictures, is told by the teacher, after a preparation in which the prior knowledge of the children is activated and the pictures are described as giving the framework of the story (step 1: the preparation phase). The teachers read the stories and discussed the 'difficult words' with the children. The teacher also explains the meaning of the target words in the lesson (step 2: semantization). After that, children do some language games, they have to participate in a dialogue or they have to perform small tasks like colouring certain objects in pictures (step 3: consolidation). Finally, the children could listen to the stories on cassette recorders in a 'listening corner' in the classroom, They could also listen to the stories in the materials developed for second language learning. In this way, the curriculum was repeated once more. It should be stressed that this type of second language teaching is in no way comparable to traditional foreign language teaching, in which the *study* of language *forms* is central. Second language teaching as in our experimental program for children up to 8 years old, relies more on acquisition of the target language in meaningful contexts and communicative tasks.

It should be stressed that it was not our intention to implement 'an ideal curriculum'. If that had been our aim, we would have selected the 5,000 words to be learned, and on the basis of those words, we would have developed a new curriculum and new course materials. Such a design was far beyond our possibilities. Our aim was to find out if it was possible substantially to speed up migrant children's acquisition of Dutch vocabulary, and the program we designed was, in our view, the best choice for that time.

4 experimental schools (57 children)
4 control schools (53 children)
Predominantly Turkish & Moroccan children
curriculum dependent pre- and posttests
curriculum independent tests on vocabulary + reading yearly, and in grade 7, too.
Questionnaires to monitor program execution

3.2 Results

With the exception of one school, the program was not performed completely, but for about 80%. The mean time spent on the program was less than prescribed (four hours a week): in grades 1 and 2 (K indergarten) and grades 3 and 4, 3.40 (hours/minutes), 3.05, 2.40 and 2.10 respectively. One school stopped all activities in the fourth year. The time spent on the program decreases in higher grades, which is certainly a consequence of the fact that the teachers give priority to other educational aims, especially learning to read and write in grades 3 and 4. It turns out to be difficult to introduce a program which is not a substitution for another program, but is considered by the teachers as additional. Repetition of stories (and target words in the stories) was important, because there is a clear relationship between learning of words and repetition. But the answers on the questionnaire showed that the picture books were not repeated as often as prescribed: two, at most three times instead of the prescribed four times (including preparatory presentation of the story on the basis of the pictures). Many times the separated Dutch L2 lessons were cancelled. Also the children generally did not listen as many times to the stories individually in the 'listening corner' as planned in the program: one or two times to each story instead of four times.

The teachers did not always have a positive attitude towards the program, which might have had a negative influence on the implementation. In the K indergarten groups some teachers found it difficult to work methodically according to a program. In their view education for young children should follow and stimulate their development in a 'natural' way; the program was too 'methodical', too much teacher-centred and resembled too much formal education for older children. As noted above, in grades 3 and 4 teachers often found it difficult to integrate the experimental program in the regular curriculum. As a result, the number of words acquired was not as high as expected.

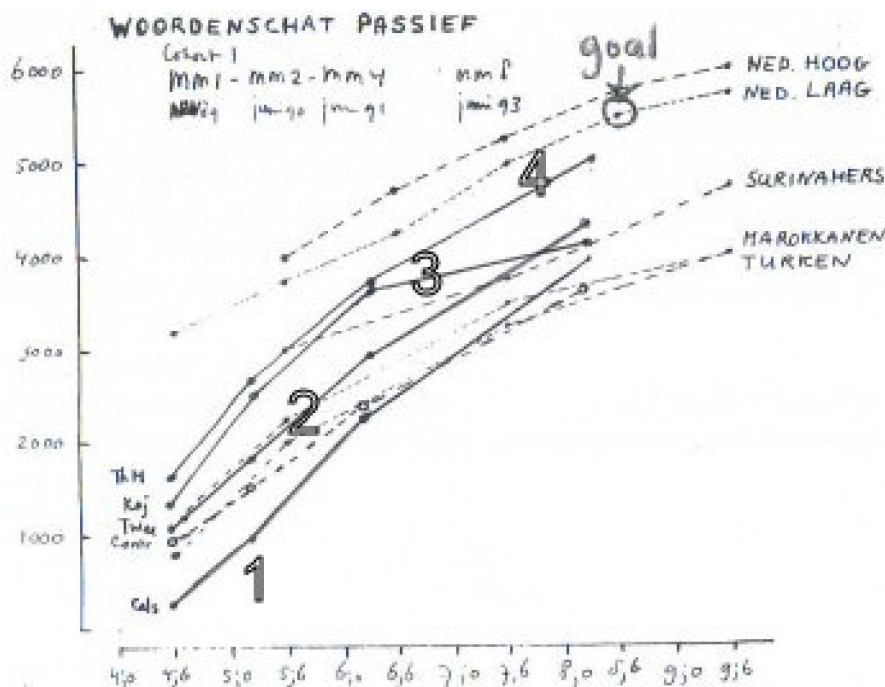


Figure 2: longitudinal scores on the curriculum-independent tests of the individual schools

However, as can be seen in the second figure, compared to the control schools, the experimental schools did much better on the curriculum independent tests. The control schools had the same scores as other Turkish and Moroccan children in the Netherlands. Of

the experimental schools, one of the schools (school 4) almost reached our goal. The school indicated with number 3, dropped all activities in the fourth year and fell back. Even school number 1, that started almost around zero, outperformed the control schools after four years. Even more promising are the results of these children in the end of grade 7. The intervention program only lasted the first four years, but probably due to the solid basic vocabulary teaching those first four years, the children almost catch up with their native Dutch classmates, as can be seen in figure 3, where the scores of experimental schools number 1 and 4 are displayed.

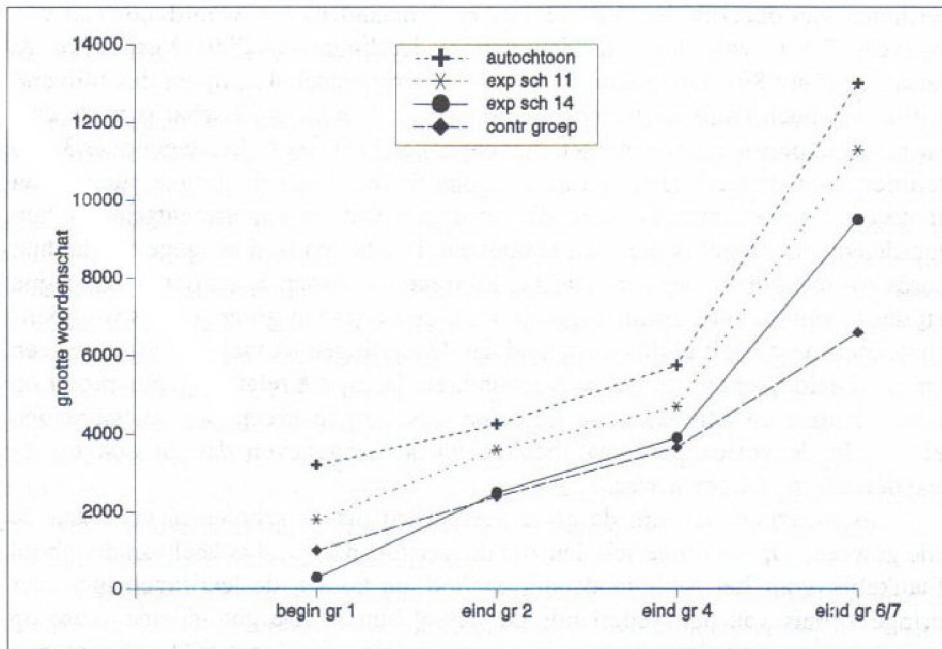


Figure 2: longitudinal scores begin grade 1, end grades 2 and 4, and end grade 7, on the curriculum-independent tests of two experimental schools, and the control schools.

Their greater vocabulary probably facilitated their learning of new words in the higher grades, due to a better comprehension of the school texts. The results of our experimental program show that it is possible to speed up the acquisition of Dutch vocabulary by migrant children in the first four grades of primary school. By the end of 4th grade they were one or two years ahead of their comparison group age peers in Dutch vocabulary. So, it is indeed possible to speed up the second language vocabulary of children in such a way that they can catch up with their native Dutch classmates in lexical skills. Such a goal is possible under the following conditions: (1) The curriculum has to offer sufficient new words in the materials, that is, at least 1,000 new words each year taught explicitly; (2) the children must have second language vocabulary instruction in small groups, at least 4 days a week, within and outside the classroom, for consolidating, repetition and evaluating.

4 Conclusions and some curriculum examples.

In the early stages of second language learning, it is absolutely necessary to organise separate classes, at least an hour or so daily, otherwise especially the younger children will drown in the language bath. They have to learn the first communicative needs as soon as possible: *what's your name, how old are you, sit down please, I don't speak Icelandic, only a little . komdu blessaður.* For the early stages of these children we developed in the Netherlands special curriculum materials, with, for example, photo stories and small dialogues:

1. naar school



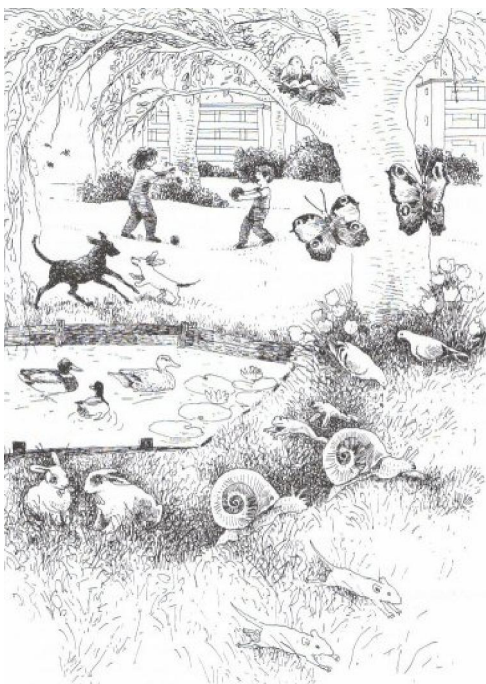
Talarðu íslensku?
Svo lítið

Do you speak Icelandic?
A little.

The various dialogues are done first with the teacher, and later on played by the children together. Remark that in these stories, there is always an easy and a more difficult role to play, thus, from day one on, a child can play its role.

From my other lecture about second language learning and teaching, and from the research project on speeding up vocabulary, it is quite clear that not only in the early stages extra lessons in the second language are necessary. These extra lessons are needed because second language learners need OTHER learning contents (such as sounds in a language that leads to different meanings of words), They need MORE contents (such as the enormous amount of words), they need OTHER EXERCISES (for instance, on morphology), and above all, they need MORE explicit input, and MORE repetition of the learning content.

An example of another type of exercises we do with second language learners is on morphology. Dutch morphology is not THAT complicated as Icelandic, (*Íslenskan er ekki einföld*) but still... ..To exercise the different types of forming the plural of nouns, illustrations of the following kind are used:



Of all these animals there is more than one, so the different forms of the plural is made very explicit to children if you work with these. See also the numerous books on counting, e.g. the famous *Anno's counting book*.

Another aspect what we learned from our experiment is that teachers with immigrant children in their class find it a difficult task to differentiate in learning content and organisation. And of course, it IS a difficult job. That is another reason more to develop special curriculum materials for immigrant children. An important essence of such materials is that children can find the solution of such exercises on their own, without using a native intuition, without having the knowledge of these forms already. Take the following examples in Dutch:

3
Maak het groene woord zo dat het in de zin past.

ophouden Mol _____ met graven.

rondkijken Hij _____ boven de grond.

eruitzien De wereld _____ anders _____!

opwerpen Wie _____ daar die aarde _____?

inhouden Mol _____ zijn adem _____.

Of kies uit:

houdt in

houdt op

kijkt rond

werpt op

ziet er uit

The children have to fill in the green word correctly in the sentences. Native Dutch children know that the *Verberuitzien* in the third sentence, appears in three separate parts, and that those parts change in order. That is their native intuition: they simply fill in: ziet - er - anders - uit. So, they just fill in the sentence. But on the right side of the exercise, the solution is offered, too. In this way, non native children can find out quite easily the correct sentence.

The lack of curriculum materials that can be done by second language learners themselves, and that can differentiate between learners of different levels of proficiency, is one of the reasons that the last years I spend much time in developing materials for the computer. In recent years, the computer has become very common in schools and books on the computer have become available. These are books presented on the computer that include such features as animations or karaoke text. These computer books allow children to interact with the story. They are very motivating, combine both audio and visual elements, and can thus contribute to literacy development, and enhance vocabulary. Recent developments also include speech recognition into the books on the computer, so children can read with the computer. Because the computer can track down the level of the child, know what mistakes the child made, for each child individual learning content can be tailored. However, developing such adaptive computer materials is extremely time consuming, and is very expensive. You can see some examples on my website www.woordwerkenanneveer.com.

In general, once the teacher is aware of the fact that second language learning is not that simple for a child, and that it does not come by itself, he can make the acquisition process easier for a child by being more explicit, by showing more, by not only pointing to an object, but also in the same time, naming the object. So, if the children during break or lunch time,

have some fruit to eat, e.g. *an apple*, do not ask: *Do you like that? Do you eat them often?* but ask explicitly naming the object: *Do you like that apple? Do you eat apples often?* In the lecture before, I gave already some other examples of explicit teaching, like in reading lessons: let the children first read the text for themselves, and let them underscore the words they do not know. And afterwards look up these wordings.

Or: Select weekly some target words and make them explicitly visible in the classroom, as in following photo:



Or help the children to understand a science text easier by writing down synonyms of difficult words in the margin of the textbook.

Underline and write down synonyms in the margin of science texts:

American democracy has usually been found wanting *deficient*
whenever the question of justice and equality for
Americans of African descent has been raised *origin*

Concluding remarks:

Extra, separate L2-lessons necessary:
OTHER learning contents (e.g., sounds)
MORE contents (e.g., words)
OTHER exercises (e.g., morphology)
MORE explicit input
MORE repetition

Differentiation in the classroom:
Offer solutions of exercises to L2-learners
'Listening corner', Adaptive ICT materials
Being explicit in word & rules & gestures