Content-based language learning in English

A model for high proficiency in English in Swedish schools?

Marika Kjellén Simes

Teachers have long wondered how a foreign language is learned effectively. It is often said that the target language is best learnt in a country where people speak it. In other words, the learner benefits from environments where the target language is used as the means of communication. In the school context, attempts can be made to recreate such natural environments by using the target language as the means of communication in the classroom to varying degrees. In some cases the target language is even used outside the foreign language classroom as the mode of instruction in one or several subjects. This method is referred to as immersion education or Content and Language Integrated Learning (CLIL). Although German, French, and Spanish are used in Swedish CLIL programmes, English stands out as the most widely used target language (Nixon 2000: 3). CLIL takes on many different forms but I have chosen to study the International Baccalaureate (henceforth IB). IB is an international school in which examination is partly carried out by external examiners. This means that IB students write quite extensively in English, possibly to a higher extent than students in most other forms of CLIL in

Sweden. Since my study focuses on student writing in English, this is an important fact that should be kept in mind.

The present paper is a brief account of the main findings in my dissertation (Kjellén Simes, 2008). In short, my research questions were the following:

- Will IB students become more proficient in English than students who only study English as a foreign language?
- Can initial proficiency be used to explain development over time?
- Can students' proficiency be explained in terms of their motivation?

Previous research

The CLIL method is popular and, according to figures from 1999, it was used in 20% of Swedish upper secondary schools (Nixon 2000: 3). It is widely used also in other parts of the world, and the reason for its popularity has to do with the use of status languages such as English as the target languages. Still, although a number of studies have been undertaken on CLIL, it is far from easy to make general claims about the benefits of CLIL. This has to do with the fact that CLIL takes on different forms in different contexts. First, the extent to which the CLIL method is used varies. In full immersion all subjects are taught in the target language, whereas in partial immersion just one or a few subjects are taught in the target language. Moreover, there is the age factor: CLIL may be used with very young children (early immersion) or be introduced to older children (late immersion). Lastly, there is the question of teacher competence in the target language. It is inevitable that standards differ not only between individual teachers, but also, due to rules and regulations, between schools and even nations.

There are three different areas which tend to be in focus in CLIL research, namely whether students actually make any target language gains, whether subject knowledge is negatively affected by the use of the target language as the mode of instruction, and whether CLIL students make native language losses because of the emphasis on the target language (see e.g. Lim Falk 2008; Dalton-Puffer 2008; Vollmer 2007). The focus of my own study was target language gains, an area where previous research has yielded quite contradictory findings. Whereas Canadian, German, and Finnish studies have all reported on positive outcomes (see e.g. Wode 1999: 26, 30; Laurén 1999: 175f), studies done in Sweden are not as favourable. For instance, Åseskog (1982), who studied the effects of the CLIL method in electrical engineering education, found that although CLIL students reached better results in English than their peers, statistical significance could not be reached. In effect that means that the differences that were found could have been due to chance. Another study by Washburn (1997) on CLIL in upper secondary

school only found positive effects in the domain of spoken fluency in English. A third study, by Sylvén (2004), suggests that gains in English might as well be attributed to the CLIL students' larger amount of spare time reading in that language. Critics of the CLIL method have pointed to factors other than the method itself as paramount in students' potential success. One of them is Josephson (2004:128, 139f), who suggests that CLIL students tend to be motivated, middle-class students, who would probably achieve well in most circumstances. A factor that further complicates the matter of target language gains is the fact that the CLIL method might be more profitable for some target language competencies than others (Dalton-Puffer 2008: 5f, 15). Hence, the use of target language literature and emphasis on oral communication in the CLIL classroom would make students better when it comes to receptive skills, vocabulary, and fluency in speaking. Competencies such as grammar and writing, on the other hand, would not be affected to the same extent. However, since IB students write extensively in the target language, it cannot automatically be concluded that this applies to them as well.

What, then, should a study about the effects of CLIL on target language proficiency focus on? One way of measuring a student's proficiency is to examine the vocabulary they use, e.g. in terms of their ratio of "rare words". This field was first explored by Moira Linnarud (1986), who studied what she called vocabulary sophistication in order to find out about learners' ability to use precise and sophisticated vocabulary. Other linguists who have developed methods for studying learners' use of "rare vocabulary" are Batia Laufer and Paul Nation. In their studies, they have focused on the frequency of words in different types of texts, arguing that proficiency is positively correlated with the use of rare or low-frequency vocabulary (see e.g. Laufer 1994: 29). This means that the more proficient the learner, the larger the proportion of low-frequency words.

Another important indicator of learners' target language proficiency is their use of verbs, since mastering verb morphology for tense, aspect and number is a demanding task.[1] Not only do learners need to master verb morphology; they also need to be able to use tense appropriately. According to Vaughan (1991: 114), the latter is indicative of learners' writing skills. It has been found that verb errors are more likely to occur when tense shifts are made (Chappell & Rodby 1983: 310f). Tense shifts as such were of interest also to Lars Sigfred Evensen (1987; Albrechtsen et al. 1991: 81f). He studied the use of tense shift in order to find out if

---

[1] Tense gives information on time orientation, e.g. *she runs*, which is in the present tense, and *she ran*, which is in the past tense. Aspect reflects whether the verb event is completed or not. The perfect aspect, e.g. *she had run*, denotes an event that took place before a specified time. The progressive aspect, e.g. *she is running*, is used for events that are in progress at a specified time. As to number, the third person singular is marked by the suffix *–(e)s* in the present tense.

they were in some way related to proficiency. In a minor pilot study (N=9)[2] of learner texts, he found a positive correlation between proficiency and the use of contextually motivated tense shifts, and a negative correlation between proficiency and contextually unmotivated tense shifts. In effect this means that the more proficient the learner, the more motivated tense shifts and the less unmotivated tense shifts are being made.

As mentioned above, critics of CLIL have pointed to motivation as a factor that is likely to guarantee some level of success for the method (Josephson 2004: 128, 139f). In a study of proficiency outcomes it is therefore necessary to take into account also the effects of learner motivation and attitudes. Among other things, it has been found that different kinds of motivation correlate in different ways with academic performance (Pintrich & de Groot 1990: 33). Learners can either be willing to learn things for their own sake because they find them enjoyable, i.e. they are intrinsically motivated, or they are extrinsically motivated, which means that they may perform activities in order to either achieve something positive or to avoid something negative. Studies have shown that school achievement is helped by intrinsic motivation (Pintrich & de Groot 1990: 33; Gottfried 1985: 642). Also, in the case of language learning, it has been suggested that learners' success rate is related to their attitudes towards target language cultures and target-language speaking people (Gardner & Lambert 1972: 3, 12ff). More specifically, positive attitudes would facilitate learning. Lastly, it has been found that the way to success can be regarded in different ways: whereas some students believe that success is based on effort, others believe that success is based on an innate ability and thus cannot be influenced by hard work. The former orientation is called *mastery orientation* and the latter *ego orientation* (Salili et al. 2001: 8). A positive correlation has been found between a number of desired academic outcomes and mastery orientation (Urdan 2001: 183).

Material

As mentioned above, my study was an attempt to find out whether students at the IB programme, as a form of CLIL, become more proficient in English than students who study English as a foreign language only. I decided to opt for a longitudinal study of written production, running for two years, as well as take into account different motivational factors. The material I studied consisted of three sets of compositions written in English by 86 students in upper secondary school. Half of them (43) were IB students, who were taught through English in all subjects, and another 43 students were selected from among 124 students at different national

---

[2] N stands for number (of informants in the study).

programmes (NP), all of them studying English as a foreign language. The first composition was written at the beginning of the first year in upper secondary school (year 10), the second composition at the beginning of the second year (year 11), and the third composition at the end of the second year (year 11). For each set of compositions the students were given a topic with an imaginary subject matter. In all, the 258 narratives collected made up a total of 126,000 running words. In addition to writing narratives, all participants filled in a questionnaire covering background factors such as motivation and attitudes towards English. At the beginning of the study the students did a diagnostic test, which enabled me to select 43 NP students with results that matched those of the 43 IB students. The students' results on the diagnostic test made it possible to divide them into three subgroups: subgroup I, with above average results, subgroup II with average results, and subgroup III with below average results. Here it is important to note that what is referred to as average results in this study is not average results on a national level, since none of the students who participated in the study had a grade below pass with distinction (VG) in year 9 in compulsory school.

Focus of the study

In order to assess the students' proficiency in English, the compositions were studied in two different ways: I identified rare words (below I refer to rare words as low-frequency vocabulary) and I also analysed the students' skills in using tense shift (see examples (1) – (3) below). Low-frequency vocabulary was identified with the help of *The Collins Cobuild Dictionary for Advanced Learners* (2001), in which words are divided into five bands according to their frequency (*Collins Cobuild*, 2001:xlii). Frequency band 5 contains a small number of words (680) which are the most frequently used words in the English language. For each band the number of words increases while their use becomes less frequent. Band 1, for instance, contains as many as 8,100 words. Together, the five bands span 14,600 words, which in fact make up 95% of the vocabulary in the texts that *Collins Cobuild* is based on, its corpus. In my study I followed Laufer (1995) and split the vocabulary of the compositions into two groups: high-frequency vocabulary (bands 4 and 5) and low-frequency vocabulary (bands 3, 2, 1, and what I refer to as band 0, which are the words with a frequency lower than the words in frequency band 1). The first group comprises 1,720 common words, such as *and*, *like*, *go*, *bridge*, and *danger*. Together, these high-frequency words constitute 75% of the vocabulary in the *Collins Cobuild* corpus (*Collins Cobuild*, 2001:xlii). The second group, low-frequency vocabulary, comprises all the remaining words, that is 25% of the vocabulary in the *Collins Cobuild* corpus. Examples of such low-frequency words are *aggressive*, *duration*, *abundant*, and *hallowed*. In order to measure the

vocabulary quality of each of the texts, the number of low-frequency lemmas[3] was divided by the total number of lemmas, arriving at the ratio of Low-Frequency Vocabulary (LFV):

$$LFV = \frac{\text{Lemmas from frequency bands } 3-0}{\text{Total number of lemmas}} \times \frac{100}{1}$$

In the compositions, the ratio of Low-Frequency Vocabulary ranged from 3.8 to 29.1.

The second approach that was taken to the student texts was that of analysing the tense shifts that occurred. When it comes to narratives, which was the text type in focus in my study, the main text tends to be kept in the past tense, whereas so-called side structures are often in other tenses. There are different types of side structures that may call for a tense shift. One of them is background, by which is meant reports on events that took place before the events in the main text, as in (1), where the simple past of the main text is accompanied by a side structure in the past perfect:

(1) She <u>went</u> to sleep after a day that **<u>had been</u>**[4] absolutely fantastic.[5]

Another reason for making a tense shift in a narrative in the simple past could be the use of dialogue, since direct speech is often in the present tense. This is illustrated in (2):

(2) Pamela <u>looked</u> at them: "You **<u>are</u>** too noisy", she <u>told</u> them.

Lastly, tense shifts also tend to be made in cases where the narrator temporarily steps out of the main text and speaks directly to the reader, so-called meta-comments and general truths. In such cases the present tense is usually used, as in (3), where the simple present is used for a meta-comment in a narrative where the main text is in the past tense:

(3) Some years ago, I **<u>think</u>** it was some time in the sixties or in the seventies, I <u>was drifting</u> around in Asia.

---

[3] Lemmas are sets of words that share the same root and belong to the same word class. For instance, *jumped* and *jumping* are both part of the same lemma, the headverb *jump*. The noun *jump*, on the other hand, forms a lemma of its own.

[4] Finite verbs are underscored. Tense shifts are underscored as well as in bold face.

[5] Examples (1) – (4) were taken from student compositions in the study.

While motivated tense shifts such as those in examples (1) – (3) help the reader navigate between the chronology of the main text and various side structures, tense shifts that occur at random, without being contextually motivated, tend to obscure the text. One such example can be seen in (4), where the narrator deviates from the simple past and, seemingly at random, uses the simple present:

(4) I <u>nodded</u> and **<u>smile</u>** back at him.

The measurement that was used for tense shift was the number of motivated tense shift divided by the total number of finite verbs: the ratio of Motivated Tense Shift (MTSh).

$$MTSh = \frac{\text{Motivated tense shifts}}{\text{Total number of finite verb phrases}} \times \frac{100}{1}$$

In the compositions that were studied, the ratio of motivated tense shift ranged from 2.3 to 31.6. In the way indicated in the formula above, I also calculated the ratio of Random Tense Shift (RTSh). In the compositions, the ratio of random tense shift ranged between 0.0 and 11.5.

Results 1: Vocabulary use

The results in the first part of my study, on low-frequency vocabulary, showed that overall, in all three sets of compositions, the IB students used a higher ratio of low-frequency vocabulary than the NP students. The difference between the groups increased over time and in the third set of compositions it was statistically significant, i.e. not due to chance (NP: 14.49, IB: 17.78, p=0.006). I also examined the use of low-frequency vocabulary for the three subgroups: subgroup I with above average results on the diagnostic test, subgroup II with average results, and subgroup III with below average results. By and large, the mean ratios of low-frequency vocabulary of the three subgroups tended to reflect the expected proficiency levels quite well, with the highest values in subgroup I, both in the NP group and the IB group. In all three compositions, the highest values were found for the IB students in subgroup I.

Over time, the results differed in an interesting way between the subgroups. While the mean values remained more or less level in two of the NP subgroups (subgroup I: from 17.03 to 16.36, subgroup II: from 15.46 to 14.93) and one of the IB subgroups (subgroup I: from 20.47 to 19.58), there was a decrease in the mean values for the NP students in subgroup III (from 13.48 to 11.89) and an increase in the mean values for the IB students in subgroups II (from 16.51 to 17.63) and III

(from 13.82 to 15.94). None of the changes over time reached statistical significance, however. Nevertheless, in subgroup III a very interesting finding was made, showing a decrease in the use of low-frequency vocabulary for the NP students and an increase for the IB students. In the last set of compositions, the difference between the NP and IB students is statistically significant (NP: 11.89, IB: 15.94, p=0.043). In fact, in the third set of compositions, the IB students in subgroup III (15.94), who started out at approximately the same low level as the NP students in subgroup III, all but reached the mean LFV of the most proficient NP students, those in subgroup I (16.36). In Figure 1 the development over time for the IB students in subgroup III is compared to that of the NP students in all three subgroups. The mean values for the IB students in subgroups I and II, all at the level of NP I or above, are not included in the figure.
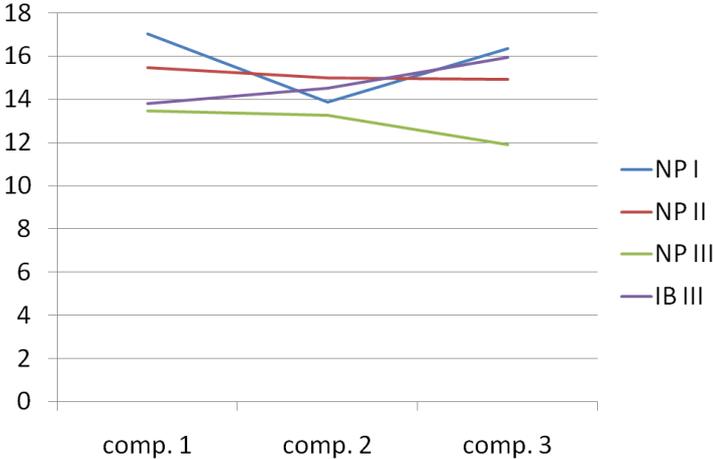

Figure 1. *Mean ratio of low-frequency vocabulary for all NP students (subgroups I-III) and the initially weakest IB students (subgroup III).*

Results 2: Tense shift

When it comes to the second part of my study, the use of tense shift, the IB and NP students initially had a similar mean ratio of motivated tense shift (NP: 13.13, IB: 12.87), but over time there was a statistically significant increase in the IB group (3.20, p=0.017) while the NP students remained at the same level. In the third set of compositions the difference between the NP students and the IB students was statistically significant (NP: 13.17, IB: 16.07, p=0.016). A comparison of the three subgroups revealed that in the NP group, the use of tense shift did not correlate as closely with the subgroups as the mean use of low-frequency vocabulary had done.

Rather, in the first composition it was quite unexpectedly the NP students in subgroup III who had the second highest figure (15.35), surpassed only by the IB students in subgroup I (16.09). However, when I looked more closely at the texts I discovered that a large part of the motivated tense shifts in the first composition of the NP students in subgroup III was due to dialogue. This suggests that tense shift due to dialogue might be less demanding than other types of motivated tense shift and therefore not as straightforwardly correlated with proficiency. That said, there were still some interesting results to be found, especially for the students in subgroup III.[6] Over time, they showed diverging results: the NP students had a decreasing ratio of motivated tense shift over time (from 15.35 to 12.32), whereas the IB students had an increase which was statistically significant (from 8.16 to 15.78, p=0.001). It should be noted that the use of tense shift due to dialogue had been reduced in the third set of compositions and appeared to a similar degree among NP and IB students in all three subgroups.
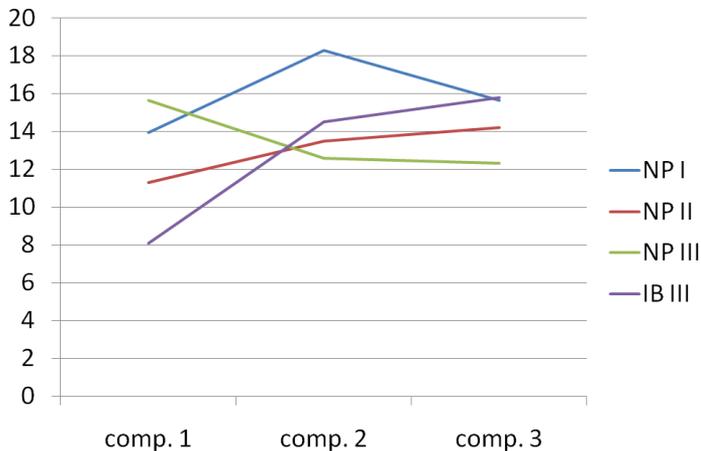

Figure 2. *Mean ratio of motivated tense shift for all NP students (subgroups I-III) and the initially weakest IB students (subgroup III).*

Figure 2 shows the development over time for the IB students in subgroup III as well as that of the NP students in all three subgroups. The mean values for the IB students in subgroups I and II, all at the level of NP I or above, are not included in the figure. Figure 2 shows clearly that in the last composition, the IB students in subgroup III (15.78) used motivated tense shift to the same degree as the most proficient NP students, namely those in subgroup I (15.64). Consequently, just as in

---

[6] For a full account of my results, see Kjellén Simes (2008)

the case of vocabulary quality, the IB students who progressed the most over time were the ones who were initially the weakest.

When it comes to random tense shift, the figures are low in all subgroups, and there is also a tendency for the mean values to decrease over time. The NP students in subgroup III, however, were the only ones who did not follow this pattern, but actually increased their ratio of random tense shift over time. In fact, in the third set of compositions the difference between the NP and IB students in subgroup III is statistically significant (NP: 3.97, IB: 1.63, p=0.022).

Results 3: Student motivation

As a complement to the investigation of the students' use of vocabulary and tense shift in the three compositions, the motivation of the participants was studied by means of a questionnaire. The students' replies revealed that motivation was indeed a factor that could help explain the differences that were found between NP and IB students' results. Statistically significant differences were found overall and in subgroup III in particular. First of all, the NP and IB students differed in their goal orientation. The IB students were mastery oriented to a higher degree than the NP students, overall and in subgroup III specifically. In effect, this means that the IB students were more convinced that academic success is the result of effort rather than innate ability (i.e. talent), something that is likely to be beneficial for their academic performance. Thirdly, the IB students overall as well as those in subgroup III reported a higher degree of intrinsic motivation than the NP students. This means that the IB students to a higher degree felt rewarded by increasing proficiency per se, which is also likely to boost the academic achievement of students. Lastly, the IB students overall as well as in subgroup III reported a higher degree of positive attitudes towards English cultures and English-speaking people than the NP students. This is also a factor that is likely to influence the academic performance of students in a positive way. Consequently, there are three different motivational factors that help explain the improvement that was found among the IB students in general and in subgroup III in particular, i.e. the initially weakest IB students.

Conclusion

The results of the two linguistic studies point in the same direction: The IB students became more proficient in written English than the NP students, both as regards vocabulary quality and the use of tense shift. The questionnaire study revealed that the IB students were better motivated than the NP students. Consequently, not only can the IB students' higher proficiency in writing be explained by the fact that they had English as the medium of instruction, they were also better motivated to learn

English. A final finding was that the biggest differences between the IB and NP students were found for the initially weakest students, those who had below average results on the diagnostic test. This was the case as regards the two linguistic features as well as motivation. The results of this study therefore suggest that, as a way of promoting improved written proficiency in a target language, the IB programme yields good results, particularly so for somewhat less proficient but motivated students.

References

Albrechtsen, D., L.S. Evensen, A.C. Lindeberg & M. Linnarud. 1991. Analysing developing discourse structure: The NORDWRITE project. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith & M. Swain (eds.), *Foreign/Second Language Pedagogy Research.* 79–91. Clevedon: Multilingual Matters,

Åseskog, T. 1982. *Att undervisa el-lära på engelska: Ett försök på el-tekninsk linje på gymnasiet.* Göteborg: Göteborgs universitet.

Chappell, V. & J. Rodby. 1983. Verb tense and ESL composition: A discourse level approach. In M. Clarke & J. Handscombe (eds.), *On Tesol '82: Pacific Perspectives on Language Learning and Teaching.* 309–320.Washington D.C.: Teachers of English to Speakers of Other Languages.

Cock, D. & H. Halvari. 2001. Motivation, performance and satisfaction at school. In Efklides et al. (eds.), 65–84.

The Collins COBUILD English Dictionary for Advanced Learners. 2001. Glasgow: HarperCollins.

Dalton-Puffer, C. 2008. Outcomes and processes in Content and Language Integrated Learning (CLIL): Current research from Europe. In: W. Delanoy & L. Volkmann (eds.), *Future Perspectives for English Language Teaching.* 1–19. Heidelberg: Carl Winter. Available at [http://www.univie.ac.at/Anglistik/Dalton/SEW07/CLIL%20research%20overview%20article.pdf].

Efklides, A., L. Kuhl & R. Sorrentino (eds.). 2001. *Trends and Prospects in Motivation Research*. Dordrecht: Kluwer Academic Publishers.

Evensen, L. S. 1987. Elevtekster og elevers tekststrategier. *Trondheimsskrifter i anvendt språkvitenskap III (TRANS - 3)*: 55–83.

Gardner, R. & W. Lambert. 1972. *Attitudes and Motivation in Second-Language Learning.* Rowley, MA: Newbury House.

Gottfried, A. E. 1985. Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology 77*(6): 631–645.

Josephson, O. 2004. *"Ju". Ifrågasatta självklarheter om svenskan, engelskan och alla andra språk i Sverige.* Stockholm: Norstedts Ordbok.

Kjellén Simes, M. 2008. *Room for Improvement? A comparative study of Swedish learners' free written English in the foreign language classroom and in immersion education.* Karlstad: Karlstad University Studies 2008:32.

Kuhl, J. 2001. A functional approach to motivation. In Efklides et al. (eds.), 239–268.

Laufer, B. 1994. The lexical profile of second language writing: Does it change over time? *RELC Journal 25* (2): 21–33.

Laufer, B. 1995. Beyond 2000: A measure of productive lexicon in a second language. In L. Eubank, L. Selinker & M. Sharwood Smith (eds.), *The Current State of Interlanguage.* Amsterdam: Benjamins, 265–272.

Laurén, C. 1999. *Språkbad: Forskning och praktik.* Vasa: Vasa University

Lim Falk, M. 2008. *Svenska i engelskspråkig skolmiljö: Ämnesrelaterat språkbruk i två gymnasieklasser.* Stockholm: Acta Universitatis Stockholmiensis.

Linnarud, M. 1986. *Lexis in Composition.* Lund: Lund Studies in English 74.

Nixon, J. 2000. *SPRINT: Content and Language Integrated Learning and Teaching in Sweden.* Stockholm: Skolverket.

Pintrich, P. & E. de Groot. 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology 82*: 33–40.

Sylvén, L. K. 2004. *Teaching in English or English Teaching: On the Effects of Content and Language Integrated Learning on Swedish Learners' Incidental Vocabulary Acquisition.* Göteborg: Acta Universitatis Gothoburgensis.

Salili, F., C. Chiu & Y. Hong. 2001. The culture and context of learning. In Salili et al. (eds.), 1–14.

Salili, F., C. Chiu & Y. Hong (eds.). 2001. *Student Motivation. The Culture and Context of Learning.* New York, N.Y.: Plenum.

Urdan, T. 2001. Contextual influences on motivation and performance: An examination of achievement goal structures. In Salili et al. (eds.), 171–201.

Vaughan, C. 1991: Holistic assessment: What goes on in the rater's mind? In L. Hamp-Lyons (ed.), *Assessing Second Language Writing in Academic Contexts*. 111–125. Norwood, NJ: Ablex.

Vollmer, H. J. 2007. Language across the curriculum: A way towards plurilingualism. In W. Martyniuk (ed.), *Towards a Common European Framework of Reference for Languages of School Education.* 177-192. Krakow: Universitas. Available at [http://www.universitas.com.pl/media/File/Fragmenty/TOWARDS/mart_2-3.pdf].

Washburn, L. 1997. *English Immersion in Sweden: A Case Study of Röllingby High School 1987–1989.* Unpublished PhD thesis. Stockholm: Stockholm University.

Wode, H. 1999. Language learning in European immersion classes. In Masih, J. (ed.). *Learning Through a Foreign Language: Models, Methods and Outcomes.* 23–36. London: CILT.