The potential of CLIL: students' use of subject-specific terminology

Angelika Rieder-Bünemann, Julia Hüttner & Ute Smit[*]

- Identification of subject-specific terms and phrases in oral CLIL data
- Quantitative and qualitative analysis of CLIL students' oral terminology use
- Insights into CLIL students' active terminology range and lexical strategies, pointing to an advantage of CLIL students but also to considerable variation between students

1. Rationale

Content and language integrated learning (CLIL) has become a widespread and well established educational practice in European schools, generating a vibrant research scene (e.g. Nikula, Dalton-Puffer & Llinares, 2013). Within this research field, studies investigating vocabulary gains through CLIL have primarily focused on students' general vocabulary, showing that overall CLIL students do better than non-CLIL learners in the area of lexical competence (Jiménez Catalán & Ruiz de Zarobe, 2009; Olsson, 2015). More recently, however, researchers have tried to conceptualize CLIL as an integrated and inseparable unit combining the areas of 'language' and 'content' (Llinares, 2015), and have consequently stressed the need to focus on the on the area where 'language' and 'content' meet. Such a link is apparent in those aspects of language, notably vocabulary, that are specific to individual (school) subjects. A number of recent studies has followed this trend, focusing on specialist language from a wider lexical perspective by analyzing stretches of subject-specific classroom discourse in CLIL (Nikula, 2012, 2015). Interestingly, though, research with a more narrow focus on CLIL students' subject-specific L2 vocabulary is notably scarce (cf. Gablasova, 2014, Heras & Lasagabaster, 2015). Importantly, given the oral nature of many classrooms, there are as yet no studies that focus specifically on oral data.

A general shortcoming of previous studies is their lack of a systematic specification of what is specialist terminology in a particular CLIL subject. Indeed, a thorough investigation of subject-specific vocabulary (SSV) gains through CLIL seems to be hampered by the fact that, as Nikula (2012, p.137) notes, "defining subject-specific language is not an easy task", since, for instance, delimiting subject-specific terms from general language vocabulary is often far

---

[*]University of Vienna, e-mail for correspondence: angelika.rieder@univie.ac.at; julia.huettner@univie.ac.at; ute.smit@univie.ac.at

from clear-cut. The definition of subject-specific terminology we opt for in this study is a broad one (see diagram 1), including both semi-technical words, i.e. words which have a common general meaning, but have also acquired a specialized meaning in a certain subject or discipline (Dudley-Evans & St. John,1998), as well as multi-word-units. This is essential as technical terminology frequently exceeds individual word boundaries (Nation 2016, p.147), for instance in the case of compounds spelled as separate words (e.g. the financial terms *balance sheet* or *exchange rate*).

An additional consideration in the CLIL context is that in some cases CLIL school subjects represent a blend of disciplines rather than narrowly corresponding to a particular scientific field. In our study, for instance, the CLIL subject comprised European economics and politics, which is a combination of two specialized disciplines. This makes some procedures suggested in the literature, like consulting specialized dictionaries (Chung & Nation, 2004), unsuitable for identifying SSV. Research on student SSV gains through CLIL thus needs to overcome both theoretical and methodological challenges.

Diagram 1: Subject-specific vocabulary in CLIL
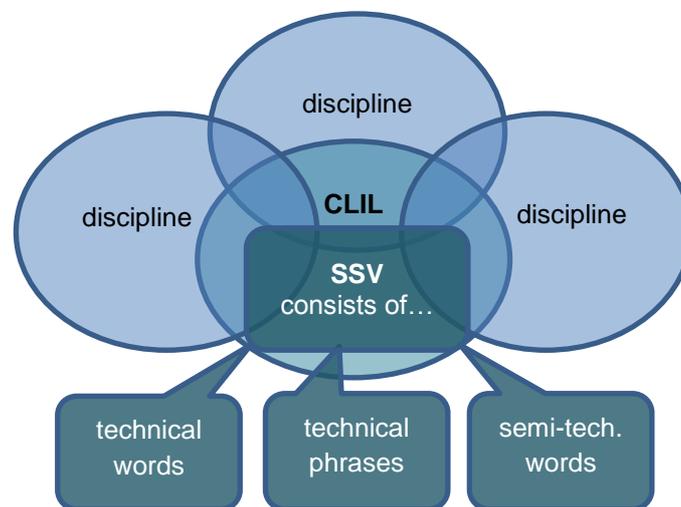
2. Research question(s)

In line with the above comments, the current study tries to identify subject-specific terminology used spontaneously by advanced CLIL students in classroom communication, in order to grasp the specific potential of CLIL to function as a link between content and language and thus go beyond providing simply more foreign language input.

The primary goal of the study clearly involves systematically quantifying student SSV knowledge as such. On the basis of the SSV identified in this process, two complementary targets are included: firstly comparing the achievements of CLIL-students and non-CLIL learners, and secondly analysing variations between individual students and SSV-related strategies observable in student communication.

Accordingly, the following research questions are asked:

RQ1

Which subject-specific terminology do CLIL students use spontaneously in oral classroom productions?

RQ2

Which of these terms exceed the estimated active vocabulary of non-CLIL students?

RQ3

What lexical variations and strategies regarding subject-specific vocabulary can be observed between CLIL-students?

3. Study description

The study was conducted in a professionally oriented college in Vienna, offering a combination of professional and pre-university academic education (HBLA – Höhere Bundeslehranstalt) with specialisations in the areas of business and tourism. The school integrates CLIL systematically as part of the curriculum in all five school years. The particular CLIL subject investigated, European economics and politics, represents a blend of two related topic areas and was taught by two CLIL-teachers, who jointly prepared the contents and materials for their lessons.

The investigation focuses on two groups of grade 12 learners (aged 17 to 18), i.e. on advanced CLIL learners who had previously experienced at least 3 years of CLIL teaching in the respective school context. English was the first foreign language they had learnt for at least seven years and, in accordance with the Austrian curriculum, up to a solid B2 level.

The dataset consists of 16 lessons that were audio- and video-recorded during one academic year, with a total of 36 learners participating in the sessions. Interaction formats during the recordings were learner-centred throughout, including partly unprepared and partly prepared activities (group work, prepared/unprepared role play, presentations). The recordings were transcribed, annotated and edited for analysis (e.g. excluding teacher turns, instances of German, etc.), with the edited data corpus amounting to 38,081 word tokens (i.e. overall number of words produced), and 2,389 word types (i.e. different words).

In order to answer the above research questions, data analysis proceeded in three steps (see diagram 2).

In the first step, a multi-methods approach was selected for identifying the subject-specific terms and phrases in the data corpus. We established a 3-phase process: Initially, a quantitative analysis employing corpus software (step 1, phase 1) was conducted in order to provide a first selection of potential SSV candidates. This was followed by a two-tier qualitative analysis carried out first by the researcher team (step 1, phase 2) and then by the CLIL-teachers as subject experts (step 1, phase 3).[†]
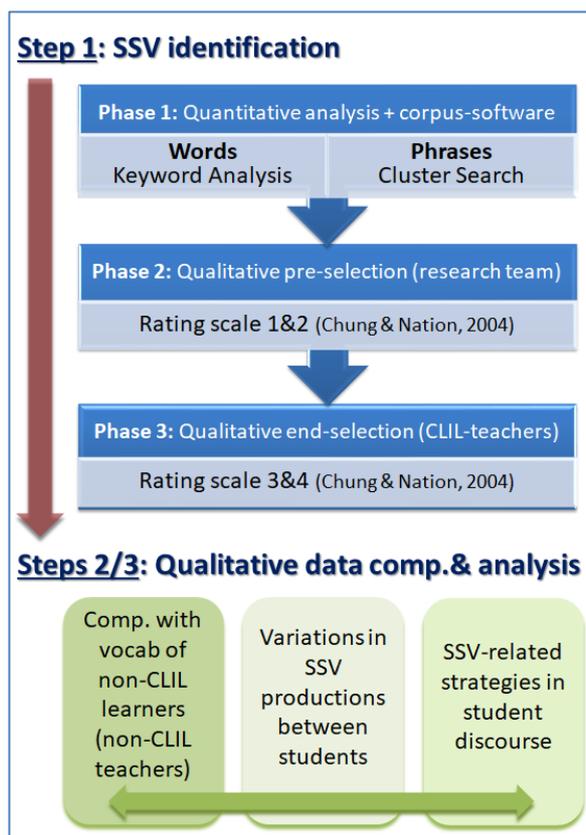

[†] A detailed description of the methodology can be found in Rieder-Bünemann, Hüttner & Smit (submitted)

On the basis of the technical terms and phrases identified (RQ1), two further steps of data analysis were carried out in order to gain more detailed insights into students' SSV use.

In step 2, the SSV identified was matched against the (estimated) active vocabulary base of non-CLIL students of the same level by obtaining estimates from four teachers at regular secondary schools (RQ2). In the third step, the data corpus was analysed qualitatively to reveal variations between individual students' SSV contributions, as well as particular strategies observed in connection with SSV use (RQ3).

Diagram 2: Research methodology

Step 1: SSV identification

Phase 1: Quantitative analysis + corpus-software

| **Words** Keyword Analysis | **Phrases** Cluster Search |

Phase 2: Qualitative pre-selection (research team)

Rating scale 1&2 (Chung & Nation, 2004)

Phase 3: Qualitative end-selection (CLIL-teachers)

Rating scale 3&4 (Chung & Nation, 2004)

Steps 2/3: Qualitative data comp.& analysis

Comp. with vocab of non-CLIL learners (non-CLIL teachers) | Variations in SSV productions between students | SSV-related strategies in student discourse

As an instrument in phase 1, we used the corpus software AntConc (Anthony, 2010). This served to check the specialized corpus against a non-specialized comparison corpus[‡] in order to retrieve potentially subject-specific one-word-items, as well as a clusters search to identify multi-word units. In phases 2 and 3, the researcher team in co-operation with CLIL-teachers carried out a four-step rating procedure (Chung & Nation, 2003) on the results of the keyword analysis.

For obtaining teacher responses concerning the estimated active SSV base of non-CLIL students, questionnaires with the lists of potential SSV items were distributed to four teachers to indicate (yes/no/comments) whether they would expect regular grade-12 language students to know the respective terms productively or not.

[‡] As a comparison corpus, we used the spoken sections of the British National Corpus (BNC; British National Corpus, 2007) and the Corpus of Contemporary American English (COCA; Davies, 2008)

4. Findings

The SSV identification in step 1 revealed a total of 70 different words (720 instances) and 52 different phrases (270 instances) in the data corpus that were classified as subject-specific by the CLIL-teachers. Various of these terms would be considered semi-technical in the sense that they occur in general language use but are used with a specialist meaning in the respective CLIL subject, i.e. European economics and politics (e.g. *balance, profit, salary, consumer, migration*). Also, the multi-word analysis showed that indeed a great number of subject-specific compounds (e.g. *austerity package, exchange rate, gender pay gap, single currency*) were only grasped through going beyond the word level; both of these findings thus confirmed the need for a broad definition of SSV.

According to the estimates of the non-CLIL teachers in step 2, over 50% of the subject-specific terms would not be expected to be part of the active vocabulary of non-CLIL pupils of the same level. A number of the semi-technical words (see above) were marked as potentially known by regular students with their general language meaning, but not with their specialist definition.

The analysis of individual student differences in step 3 showed substantial variations between students, both with regard to individual student participation in the recorded sessions, ranging from a minimum of 63 words (student C1) to a maximum of 4563 words (student C2), as well as with regard to the number of subject-specific terms produced by individual students in cases where their degree of participation was similar (e.g. for student with overall contributions of approximately 1100 to 1200 words, individual SSV productions varied between 6 SSV types and 18 SSV types).

Partly, the variations noted above are due to the fragmentary nature of the data set. Clearly, recordings of students' spontaneous classroom productions alone only grasp a fraction of what they (might) know. In order to paint a more comprehensive picture of student's SSV knowledge, complementary data such as systematic vocabulary tests focusing on relevant SSV would be required. Still, the notable variations in SSV use between students with comparable discourse contributions provide revealing insights.

Furthermore, complementary qualitative data analysis shows that instances where students produce subject-specific words or phrases are often mere repetitions of a term previously introduced by another student, so that these SSV productions cannot automatically be counted as active vocabulary knowledge (see example 1). This shows that purely quantitative results from frequency counts have to be handled with care.

Example 1:

S4:	but on the other side it's also good that we have that .. that **subsidiarity**
A1:	**subsidiarity**

At the same time, however, these repetitions can also be seen as targeted learning strategies, in the sense that subject-specific vocabulary might be repeated deliberately in order to facilitate retention. This conscious effort in relation with SSV is also underscored by the observation that the CLIL students generally show an awareness of the importance and special status of SSV within the CLIL subject (see example 2).

Example 2:

J2:	okay so a disadvantages a disadvantage is uh a disadvantage with a common currency is that **false money** wie sagt man das keine ahnung wir haben irgendwann amal gelernt was geldwäsche heißt [*how do you say that, no idea, at some point we learned what money laundering is*]
B1:	**washing**
J2:	na warte das hamma auch gelernt [*wait, we learned that, too*]
B1:	**money washing** @

As is obvious from the above examples, terminology searches do not always lead to the correct expression (which in example 2 would have been *money laundering*), but from a cognitive perspective, these instances clearly show that the CLIL students employ targeted communicative strategies to specify unknown or unclear SSV items.

4. Conclusion

In sum, the current study sheds light on various aspects of researching subject-specific vocabulary in CLIL:

- Any definition of subject-specific terminology needs to go beyond the single word to include phrases, and also needs to take into account semi-technical terms.
- The methodology chosen for identifying SSV is a viable method for extracting subject-specific terms from oral classroom data.
- CLIL students seem to successfully use a substantial number of subject-specific words and phrases spontaneously during classroom discourse which go beyond the estimated productive knowledge of students in mainstream schooling. This underlines the unique advantage that CLIL can offer as a content-language interface which exceeds the potential learning gains through offering additional general language teaching.
- At the same time, considerable variation was observed between students, indicating that the level individual students reach with regard to SSV knowledge in a CLIL setting may vary substantially.
- Regardless of this variation, students seem to be very much aware of the importance and the special status of SSV, which is shown by various learning and communication strategies observable in the data.

Thus, despite some methodology-dependent limitations, this study offers first classroom discourse based results that indicate the potential of CLIL for upper secondary level students and their mastery of subject-specific vocabulary.

References

Anthony, L. (2010). *AntConc.* [Computer Software]. Tokyo, Japan: Waseda University.

British National Corpus, version 3 (BNC XML Edition). (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Retrieved from http://www.natcorp.ox.ac.uk/, 03.05.2017.

Chung, T. M. & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language* 15(2), 103-16.

Chung, T. M. & Nation, P. (2004). Identifying technical vocabulary. *System* 32(2), 251-263.

Davies, M. (2008). The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Retrieved from http://corpus.byu.edu/coca/, 03.05.2017.

Dudley-Evans, T. & Saint John, M. (1998). *Developments in ESP: A multi-disciplinary approach.* Cambridge: CUP.

Gablasova, D. (2014). Learning and retaining specialized vocabulary from textbook reading: Comparison of learning outcomes through L1 and L2. *Modern Language Journal* 98(4), 976–991.

Heras, A. A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70-88.

Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL vs. non-CLIL instruction. In Y. Ruiz de Zarobe, & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe,* (pp.81-92). Bristol: Channel View Publications.

Llinares, A. (2015). Integration in CLIL: A proposal to inform research and successful pedagogy. *Language, Culture and Curriculum*, 28(1), 58–73.

Nation, P. (2016)*. Making and using word lists for language learning and testing.* Amsterdam: John Benjamins.

Nikula, T. (2012). On the role of peer discussions in the learning of subject-specific language use in CLIL. In E. Alcón Soler & M. P. Safont-Jordà (Eds.). *Discourse and language learning across L2 instructional settings,* (pp.133–153). Amsterdam: John Benjamins.

Nikula, T. (2015). Hands-on tasks in CLIL science classrooms as sites for subject-specific language use and learning. *System* 54, 14-27.

Nikula, T, Dalton-Puffer, C. & Llinares, A. (2013). CLIL classroom discourse: Research from Europe. *Journal of Immersion and Content-Based Language Education* 1 (1), 70–100.

Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna Språk* 109(2), 51-74.

Rieder-Bünemann, A., Hüttner, J. & Smit, U. (submitted). Capturing technical terms in spoken CLIL: A holistic method for identifying subject-specific vocabulary. *Journal of Immersion and Content-Based Language Education.*