

L2 Interactions in Heterogeneous Learner Groups during Content and Language Integrated Learning: The Experience of *Rail.lexis* and beyond

Julia Edeleva, Martin Neef, Jiaming Liu, Martin Scheidt

TU Braunschweig

y.edeleva, martin.neef, jiaming.liu, m.scheidt@tu-braunschweig.de

Abstract

Content and language integrated learning is considered a powerful tool to promote inclusion in educational settings of learners for whom the language of instruction is their additional language. Language-related difficulties of those learners have been claimed detrimental for attaining personal educational goals. Academic language places increased cognitive demands on the learning process in general due to 1) its internal complexity; 2) L2 speakers' lower proficiency; 3) their disadvantage in terms of real-time processing. Facilitators are, therefore, encouraged to integrate interactional CLIL-elements (e.g., scaffolding) during content instruction that provide the necessary pedagogical support for better understanding of disciplinary concepts and their interrelation. In the current contribution, we present the concept and first results of *Rail.lexis*, a collaborative project of the Department of German Studies and the Department of Railway Engineering at TU Braunschweig. We present and discuss several conversational arrangements (e.g., word guessing games, a differential task matrix) that were designed to engage the learners of heterogeneous linguistic backgrounds in meaningful interactions in subject-specific classes. Subject-specific tasks are gradient regarding their cognitive complexity and the background knowledge required to solve them. Therefore, the linguistic repertoire required to negotiate different task types is also differential to ensure the participation of linguistically diverse students in language-enhanced classroom interactions.

1 Introduction

A standardized language test is an essential requirement to be admitted to foreign-language study programs for learners whose preferred language deviates from the language of instruction. The language requirements for university degree studies remain quite demanding. In German universities, at least a B2 level of CEFR ([Council of Europe, 2020](#)) is required for most study programs. A B2-level language learner is described as able to “obtain information, ideas, and opinions from highly specialized sources within their field. S/he can follow the essentials of lectures, talks, and reports and other forms of academic/professional presentation which are propositionally and linguistically complex. S/he can produce clear, detailed text on a variety of subjects related to their field of interest, synthesizing and evaluating information and arguments from a number of sources” ([Council of Europe, 2020](#)). While the number of foreign applicants who fulfil the admission criteria for German-speaking study programs has been continuously increasing since 1980 ([Statistisches Bundesamt, 2023](#)), around 20 to 40% of the enrolled foreign applicants fail to attain academic goals and quit prematurely without obtaining a degree ([Heublein et al., 2020](#)). The construct of academic success is highly subjective and grounded both in individual factors and in cultural, social, and institutional integration. Language skills constitute an individual's personal profile and are subject to change over time. Wisniewski

and colleagues (2022) point out that the actual language proficiency level of foreign students in German barely reached B2.2 when they were screened at the beginning of the study programme. With around 40% of students, the language skills fail to progress beyond their initial level during their degree studies. Simultaneously, language-related difficulties have been claimed detrimental for attaining personal educational goals. Trenkic & Warmington (2019) studied the language skills of Chinese students in relation to their content-specific academic achievement in sociology. The researchers observed that academic achievement is strongly predicted by the higher- and the lower-level linguistic processing alike. Both letter naming fluency and more complex skills such as reading comprehension accuracy were equally prognostic of academic outcomes.

Regarding the linguistic integration of L2 students, numerous preparatory and in-study courses are provided by language centers at universities. They are commonly delivered as a one-size-fit-all offer and focused on developing general academic literacy and targeted strategies for taking standardized language tests. Subject-specific vocabulary comprising basic terms and collocations for a particular field of study lies outside the language course curriculum. At the same time, it constitutes the basis of successful functioning in a technical language and ensures stable academic progress in more advanced subject-specific modules. Therefore, it appears critical to identify and implement pedagogical activities to support language growth of those learners beyond passing a standardised admission test. Automated dialogue systems, or collaborative conversational agents, might be practical in self-directed learning settings (de Araujo et al., 2024). Yet the potential of collaborative conversational agents to sustain productive academic talk has been mostly restricted to operationalising talk moves that represent selected academic functions such as recapping, rephrasing, agreeing or disagreeing (de Araujo et al., 2024). The cognitive demands of the subject-specific task itself have been barely addressed to define the intervention type to be provided. Previous studies (de Araujo et al., 2024; Valle Torre et al., 2023) observed that authentic dialogue patterns may provide reliable estimations of how the learners handle the academic functions for productive discussions. The current

contribution elaborates these findings adding a further dimension of cognitive task complexity. We report initial findings regarding interaction patterns delivered by students across different conversational arrangements. We argue that cognitive task complexity should be factored in to optimise the performance of collaborative conversational agents.

2 Language-Enhancing Tools for Content Instruction in Heterogeneous Learner Groups

Supportive methodological tools are generally beneficial to generate language-enhanced instructional settings in content-driven classes which become more inclusive for L2 learners. Yet, the design of appropriate study materials remains one of the major challenges in the implementation of content and language integrated learning (CLIL; Bouvellan, 2014). In selected CLIL-frameworks (the 4Cs Framework, Coyle, 1999, the Quadrant Matrix, Cummins, 1981; the European Framework for Teacher Education, Marsh et al., 2011; the CLIL Pyramid, Meyer, 2010) as well as independent position papers and practical guidelines (Mehisto, 2012; Morton, 2013; San Isidro et al., 2020), content represents personalized knowledge that is constructed and re-constructed in learning interaction. Further, knowledge accrual occurs in resolving cognitively complex tasks which involve higher-order cognitive processes such as thinking and reasoning. Thus, cognitive functioning represents a separate domain that undergoes gradual development in a CLIL-enhanced classroom. Importantly, it is concurrent with specific language demands required to verbalize one's reasoning patterns. Thus, the linguistic, and cognitive alignment is an important prerequisite to instantiate language-enhanced interactions during content instruction. Though context-embedded interaction is fundamental for learning to take place, researchers document low proportions of specific academic functions (e.g., hypothesizing or prediction) in classroom interactions (Dalton-Puffer, 2007).

In the current contribution, we present the concept and first results of *Rail.lexis*, a collaborative project of the Department of German Studies and the Department of Railway Engineering at TU Braunschweig (Germany). One of the goals is to produce cognitively appropriate

instructional materials to instantiate meaningful classroom interactions that are conducive of knowledge (re-)construction. We also probe selected conversational arrangements as to their didactic potential to promote peer-to-peer interactions in linguistically diverse learner groups. Those goals are further developed in a follow-up project *DaF-Z mit Nachhaltigkeit (Sustainability in German as a Second and Foreign Language Teaching)* whose main aim is to make teacher professionalization more diversity-sensitive by providing sustainable and technology-enhanced language learning arrangements across contexts.

In the presentation, we first survey some linguistic and psychological preliminaries that generate a comprehensive framework to assess the appropriateness of classroom interactions for targeted language-enhanced content instruction. We present two different types of conversational arrangements where the interactions of students of diverse linguistic backgrounds around basic terms were instantiated during content instruction. The first one is a word guessing game and the second is a differential task matrix which incorporates activities promoting language and cognitive growth. Based on the preliminary results which will have been evaluated by the conference date we will contrast the interaction patterns of the learners in various types of word guessing games. Finally, we discuss how the interaction patterns of linguistically heterogeneous learners in subject-specific tasks of varying cognitive complexity can inform the design of collaborative conversational agents.

3 Linguistic and Psychological Preliminaries

The dialogue constitutes a core unit of language use. It represents a flexible, yet conversation-sustaining alternation between the speaker and the hearer who are cooperating in a goal-oriented way. Engaging in fruitful and high-quality peer interactions is positively associated with learning outcomes in various contexts (Asterhan & Schwarz, 2016; Stahl et al., 2014). Several design-based research attempts have emerged to identify the characteristics of effective collaborative behavior in dialogue-based activities. Thus, academic productive talk (APT, Michaels & O'Connor, 2015; Resnick et al., 2010) operates on the following accountability principles:

- The learners should build on and develop one another's ideas to sustain a goal-oriented interaction.
- The validity of the contributions should be secured via available reference materials or direct evidence.
- The learners should logically connect their arguments, evaluate their cohesion, and draw inferences.

While the learners' reasoning is prioritized over correctness, those classroom discourse frameworks do not give sufficient attention to studying specific reasoning patterns as a gateway to explicating mental models.

In Edeleva et al. (2024), we follow the procedures of cognitive task analysis (CTA, Klein & Militello, 2001). CTA is applied to work related tasks (e.g., generating a weather forecast or detecting an infection in a neonate) and represents a collection of methods to research, identify and represent the mental processes that evolve during task performance. CTA tasks are grounded in an extensive knowledge base and require complex inferences and judgements in a complex uncertain real-time environment. Proficiency-related differences of task performers will be stipulated in the strategies that they adopt to optimize their behaviour. Those differences are grounded in subject-related knowledge structures and mental models that underlie decision making and might be more elaborate in experts compared to novices.

Simultaneously, socio-cultural approaches to language ("Five Graces Group", 2009) stipulate that it is grounded in a specific socio-cultural context. Its emergence is concurrent with knowledge accrual. As knowledge is co-constructed, the learners' linguistic repertoire replenishes and becomes more diversified. They acquire disciplinary concepts as basic terms and negotiate the relations between them through academic discourse functions. Explicit reasoning in CTA-fashioned tasks will provide a window into the mental processes of students and how they employ language as a vehicle to re-organize their knowledge patterns. Those processes will be conducive of language growth proper.

4 Conversational Arrangements

4.1. Word Guessing Games

In Edeleva et al. (2024), we contrasted scaffolding patterns of L2 German students in

Railway Engineering in a word guessing game and a content-specific problem-solving activity. The word guessing game resembled the well-known Tabu game. The participants took turns to explain selected basic terms pertaining to the field of railway operation, albeit some intuitive explanation routes (the use of word parts or word forms, abbreviations, gestures, imitations) were eliminated by the game mechanics. The students' guessing attempts triggered meaning negotiation through linguistic adaptation. Yet, the types of scaffolds that emerged in the word guessing game differed proportionately from the strategies that emerged in a common problem-solving task. The students were less inclined to use functional and relational descriptions and embed the terms into a relevant situational context (e.g., defining initial states for a particular signal positioning). Instead, they resorted to more general factual characteristics that are contained in textbook definitions. When their initial explanation routes failed, they made use of more available prompts such as everyday meanings of the terms (e.g., *Durchrutschweg*/ *Eng. overlap* and *rutschen*/ *Eng. slip*). We conclude that pedagogical interventions should be equipped with supportive materials to gear the students' explanations in a more targeted way (cf. [Vollmer, 2008](#) for similar findings).

In a follow-up study, we proceed by surveying and comparing peer-to-peer interactions in two alternative game designs. The first game is a version of a well-known "Who is the Spy?" game. The action takes place in a city where all the "citizens" receive one and the same term, the "spy" receives a related word. The "blanco" receives a blanco card without any word. Game players take turns to describe the target term. In giving their hints, they should prevent the spy from guessing the target word. After each round, the participants vote as to who they suspect to be the spy. A still other version of a word guessing game is an adaptation of "What is on my head?" where players cooperate in their word guessing attempts. The third player in a group can provide hints to steer the guessing attempts.

4.2. Adaptive Subject-Specific Tasks

We now present the differential task matrix (DTM, Figure 1) as a didactic tool that aligns the cognitive and the linguistic domain through academic discourse functions. The matrix follows

Welche Gleisführung ist jeweils der Durchrutschweg? 	Welche weitere Durchrutschwege benötigt? 	Wie wird in ETCS Level 2 der Durchrutschweg signalisiert?
Wie wird Warteisenbahnung mit Hilfe beschildert sein? 	Warum sind diese als Durchrutschwege in Betrachtung nicht geeignet? 	Welche sind die Gleisführungen "weniger geeignet"?
Wie wird bei einer weichen Gleisführung der Durchrutschweg signalisiert? 	Wie handelt man bei unebenen Gleisführungen mit Warteisenbahnung? 	Was passiert es bei Durchrutschwegen von C2 (-) in Fahrtrichtungen mit den Signalen der Warteisenbahnung, damit die Warteisenbahnung nicht mehr passiert?

Figure 1: Example of a DTM on the topic *Overlap*. A photocopyable version can be found at <https://zenodo.org/records/7689889>

the cognitive component of Bloom's taxonomy of learning ([Bloom, 1956](#)) that is originally comprised of six levels: Knowledge, Comprehension, Application, Analysis, Synthesis, Evaluation. Originally, Bloom's taxonomy was developed to rank educational objectives based on the complexity of skills and understanding. It builds on the idea that learning is ongoing and builds on prior knowledge and skills. The taxonomy ranks respective thinking skills from least to most complex along the learning trajectory. Accordingly, learning goals can be defined and learning activities can be designed. A revised taxonomy was introduced by [Anderson & Krathwohl, 2001](#). While the original typology represents a hierarchy of educational goals, the revised typology aligns instruction, testing and assessment. It groups the cognitive operators into four knowledge dimensions:

- Knowledge of essential facts, terminology and further details that are basic to a particular discipline (factual knowledge).
- Knowledge of classification principles, theories, models, or structures pertinent to a particular discipline (conceptual knowledge).
- Knowledge of procedures and methodologies that allows the learners to modify something within a particular discipline (procedural knowledge).
- Strategic or reflective knowledge as to how to solve complex problems and tasks (metacognitive knowledge).

We re-defined the taxonomic relations between the knowledge dimensions and the cognitive operators to accommodate subject-specific instructional expectations and strategies and manipulated the level of thematic abstraction from individual facts and terms over structures and procedures to complex models (Greiner et al., 2019). The adjustments yielded a three-by-three matrix. Each cell contains individual tasks of varying complexity from A1 to C3. The learners have to negotiate specific problems that are framed to trigger recognition, manipulation or explication of disciplinary phenomena or states.

5 Initial Findings and Future Directions

The results of the first round of implementation (16 Civil Engineering students) show that the DTM was appraised by the students due to its practical utility for self-assessment and tracking of one's learning progress. We were interested in how the students navigate through the matrix. The learners had not been preliminarily advised about task-related differences in complexity. We observed that the hierarchy of difficulty implied in the matrix in terms of cognitive complexity and the degree of abstraction is perceived differently. In part, the preferred order in which the problems were solved was determined by their knowledge of the topic as well as subject-related competences and experiences. More expert students followed the reading direction from left to right to pick out the problem that they will be solving next. By contrast, the students with reduced subject-related proficiency were equally challenged by every problem regardless of its implied complexity level. Further on, the number of terms utilised by different learner dyads ranged from 23 to 136. The use of terms might be regarded as an approximation of the learners' available knowledge base. Thus, the DTM appears to elicit interaction patterns that discriminate between the students at different stages along their learning trajectory.

The DTM could benefit from multiple test runs and feedback loops from various learner groups to optimize relational item difficulty and achieve greater comprehensibility regarding the order in which the learners progress through the matrix. Though the primary goal of the matrix was to enhance learner interactions in content-enriched environments, particularly L2 learners whose language skills were compromised failed to

engage in meaningful interactions. Those learners could be supported by additional material scaffolds (De Backer et al., 2016; Martin et al., 2019) in form of task-related prefabricated chunks, linking phrases and expressions to verbalize specific academic discourse functions. Linguistic scaffolds can also be integrated as part of the conversational agents' discourse repertoire to enhance L2 learners' linguistic development. Thus, the study delivers further compelling evidence on how technology-enhanced collaborative learning should be designed to ensure academically productive talk across different conversational arrangements. Since the DTM follows the revised taxonomy which aligns learning and assessment, the interaction patterns can also be used to develop technology-enhanced assessment tools and procedures.

Acknowledgments

The project *Rail.lexis* was financially supported from 2022 till 2023 by the Lower Saxony Ministry of Science and Culture under the "Innovation plus" funding scheme. Ongoing research is being carried out at TU Braunschweig within the project *DaF-Z mit Nachhaltigkeit* which has been granted to the first author by the Innovation in Higher Education funding agency under the Open Space funding scheme (Grant number FR-546/2023).

References

- Lori W. Anderson, and David R. Krathwohl (Eds.). 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Boston, MA: Allyn & Bacon (Pearson Education Group).
- Christa S.C. Asterhan, and Baruch B. Schwarz. 2016. Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist*, 51(2): 164-187. <https://doi.org/10.1080/00461520.2016.1155458>
- Benjamin S. Bloom (Ed.). 1956. *Taxonomy of Educational Objectives: Handbook I: The Cognitive Domain*. New York: David McKay Co Inc..
- Eveliina Bouvellan, E. 2014. *Teachers' beliefs about learning and language as reflected in their views of teaching materials for content and language integrated learning (CLIL)*. Juväskylä Studies of Humanities. [Unpublished PhD manuscript, Juväskylä University].

- Council of Europe. 2020. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*. Strasbourg: Council of Europe Publishing.
- Adelson de Araujo, Pantelis M. Papadopoulos, Susan McKenney, and Ton de Jong. 2024. A learning analytics-based collaborative conversational agent to foster productive dialogue in inquiry learning. *Journal of Computer Assisted Learning*, 2024:1-15. <https://doi.org/10.1111/jcal.13007>
- Do Coyle. 1999. Theory and planning for effective classrooms: Supporting students in content and language integrated learning contexts. In John Masih (Ed.). *Learning Through a Foreign Language: Models, Methods, Outcomes*. Lancaster: CILT, pages 46-62.
- James Cummins. 1981. The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.). *Schooling and Language Minority Students: A Theoretical Framework*. National Dissemination and Assessment Center, pages 3-49.
- Christiane Dalton-Puffer. 2007. *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.20>
- Liesje De Backer, Hilde Van Keer, and Martin Valcke. 2016. Eliciting reciprocal peer-tutoring groups' metacognitive regulation through structuring and problematizing scaffolds. *The Journal of Experimental Education*, 84 (4): 804-828. <https://doi.org/10.1080/00220973.2015.1134419>
- Julia Edeleva, Martin Neef, Martin Scheidt, Gina Do Manh, and Yu Xu. 2024. Using a word guessing board game to elicit peer-scaffolds in content and language settings in Germany. In Alexiou Thomai, and Athanasios Karasimos (Eds.). *Board Games in the CLIL Classroom: New Trends in Content and Language Integrated Learning*. Berlin, Boston: De Gruyter Mouton, pages 243-262.
- Ulrich Heublein, Johanna Richter, and Robert Schmelzer. 2020. *Die Entwicklung der Studienabbruchquoten in Deutschland*. (DZHW Brief 3|2020). Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) / *The development of dropout rates in Germany* (DZHW Report 3|2020). Hannover: German Centre for Higher Education and Science Research. https://doi.org/10.34878/2020.03.dzhw_brief
- Five Graces Group. 2009. Language is a complex adaptive system. Position paper. *Language Learning*, 59 (s1): 1-26.
- Franziska Greiner, Nicole Kämpfe, Dorit Weber-Liel, Bärbel Kracke, and Julia Dietrich. 2019. Flexibles Lernen in der Hochschule mit Digitalen Differenzierungsmatrizen. *Zeitschrift für Hochschulentwicklung // Arranging flexible learning in higher education via digital differential matrices*. *Journal of Higher Education Development*, 14(3): 287-302. <https://doi.org/10.3217/zfhe-14-03/17>
- Garry Klein, and Laura Militello. 2001. Some guidelines for conducting a cognitive task analysis. In Eduardo Salas (Ed.). *Advances in Human Performance and Cognitive Engineering Research. Volume 1*. Leeds: Emerald Publishing Ltd, pages 163-199.
- David Marsh, Peeter Mehisto, Dieter Wolff, and María Jesús Frigols Martín, M.J. 2011. *European framework for CLIL teacher education: Framework for the professional development of CLIL teachers*. European Centre for Modern Languages.
- Nicole D. Martin, Catherine Dornfeld Tissenbaum, Dana Gnesdilow, and Sadhana Puntambekar. 2019. Fading distributed scaffolds: The importance of complementarity between teacher and material scaffolds. *Instructional Science* 47(1): 69–98. <https://doi.org/10.1007/s11251-018-9474-0>
- Peeter Mehisto. 2012. *Criteria for producing CLIL learning material*. *Encuentro*, 21: 15-33.
- Oliver Meyer. 2010. Towards quality-CLIL: successful planning and teaching strategies. *Pulso. Revista de educación*, 33: 11-29. <https://doi.org/10.58265/pulso.5002>
- Sarah Michaels, and Catherine O'Connor. 2015. Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion. In Lauren B. Resnick, Christa S.C. Asterhan, Sherice N. Clarke (Eds.). *Socializing Intelligence through Academic Talk and Dialogue*. Washington DC: American Educational Research Association, pages 333-347.
- Tom Morton. 2013. Critically evaluating materials for CLIL: Practitioners' practices and perspectives. In John Gray (Ed.). *Critical perspectives on language teaching materials*. London: Palgrave Macmillan, pages 111-136. https://doi.org/10.1057/9781137384263_6
- Lauren B. Resnick, Sarah Michaels, and Catherine O'Connor. 2010. How (well structured) talk builds the mind. In Robert J. Sternberg, and David D. Preiss (Eds.). *From genes to context: new discoveries about learning from educational research and their applications*. New York: Springer, pages 163-194.
- Xabier San Isidro-Smith, Do Coyle, and Sulushash I. Kerimkulova. 2020. *CLIL classroom practices in multilingual education in Kazakhstan*. Astana:

Nazarbayev University Graduate School of Education.

Gerry Stahl, Ulrike Cress, Sten Ludvigsen, and Nancy Law. 2014. Dialogic foundations of CSCL. *International Journal of Computer-Supported Collaborative Learning*, 9(2): 117-125. <https://doi.org/10.1007/s11412-014-9194-7>

Statistisches Bundesamt. 2023. Statistical report on the number of foreign students in Germany by type of educational institution and country of origin.

Danijela Trenkic, and Meeha Warmington. 2019. Language and literacy skills of home and international university students: How different are they, and does it matter? *Bilingualism: Language and Cognition*, 22 (2): 349-365. <https://doi.org/10.1017/S136672891700075X>

Manuel Valle Torre, Catharine Oertel, and Marcus Specht. 2023. The sequence matters in learning - a systematic literature review. *LAK '24: Proceedings of the 14th Learning Analytics and Knowledge Conference*, pages 263-272. <https://doi.org/10.1145/3636555.3636880>

Helmut Johannes Vollmer. 2008. Constructing tasks for content and language integrated learning and assessment. In Johannes Eckerth, and Sabine Siekmann (Eds.). *Task-Based Language Learning and Teaching. Theoretical, Methodological, and Pedagogical Perspectives*. Frankfurt/M.: Peter Lang, pages 227-290.

Katrin Wisniewski, Wolfgang Lenhard, Leonore Spiegel, and Jupp Möhring (Eds.). 2022. *Sprache und Studienerfolg bei Bildungsausländerinnen und Bildungsausländern /Language and Study Success in German-as-an-Additional-Language Students*. Münster: Waxmann.