Tectonics and Technical Language
Development of and Research on a Language-Sensitive Learning Design on Plate Tectonics

Santina Wey, University of Erlangen-Nuremberg, Germany
Dietmar Gölitz, University of Erlangen-Nuremberg, Germany
Jan C. Schubert, University of Erlangen-Nuremberg, Germany

The European Conference on Language Learning 2020
Official Conference Proceedings

Abstract
Results of international studies such as PISA and PIRLS indicate that there is a substantial correlation between language competencies, socio-economic background, and knowledge. At secondary schools, language requirements increase due to a higher level of technical language. As a result, access to subject-related learning becomes more difficult, especially for pupils with low socio-economic status or a migration biography. To give these groups of pupils equal educational opportunities, a learning design for language-sensitive teaching in the Geography classroom was developed. Research and development aspects were intertwined within the framework of design-based research. Firstly, design-frameworks were developed based on the current state of research and then transferred to a six-hour teaching unit on plate tectonics for middle school pupils, which was accompanied by research. The results of the first cycle (N=135) led to a re-design and a second cycle (N=185). To evaluate the language-sensitive unit empirically, the treatment was contrasted with a control group in a pre-post-follow-up design. Both groups covered the same content in the same period. Knowledge, technical language, and motivational aspects were measured. Learning gain scores were evaluated for validity and differences between the two groups. The results show that the experimental group has significantly higher learning gains (Cohen's d post-pre = .51) regarding geographical knowledge and technical language.

Keywords: Technical Language, School Academic Language, Subject-Related Language Learning, Geography Teaching, Language-Sensitive Lessons, Design-Based Research

1 Introduction

A glance at the current state of research suggests that language competence is a central premise for educational success. Following the so-called "PISA shock" (2000) and the comparatively weak results of students from some industrial countries, a series of additional language support measures such as extra language lessons were introduced (Mullis et al., 2016a, 2016b; OECD, 2019). What remains despite these efforts are difficulties in understanding the subjects. Educational research has explored the possible reasons. It has shown that each subject has specific language requirements and that these are aggravating factors in learning. Particularly in mathematics education, but also in physics, biology, and history, the correlation between language and subject competence has been widely documented (Handro, 2018; Härtig, 2010; Höttecke et al., 2017; Paetsch et al., 2015; Prediger & Hein, 2017; Schmiemann, 2011). As a further consequence, scientists and teachers developed and implemented language promotion measures in the form of language-sensitive subject teaching in mathematics, physics, and some social sciences. Studies show positive effects of language-sensitive teaching units on knowledge and technical language competences of children (Prediger & Zindel, 2017; Scheuer et al., 2010; Schmölzer-Eibinger, 2013; Wessel, 2015). Language-sensitive geography teaching understands language requirements in geography lessons as a learning object and as key to getting access to geographical knowledge. Criteria for language-sensitive geography teaching and empirical studies on its effectiveness, however, are lacking. The presented research project focuses on this research gap and explores two research questions:

(1) What are the design criteria of language-sensitive geography teaching?

(2) How effective is language-sensitive geography teaching compared to geography teaching without language-sensitivity regarding the acquisition of geographical knowledge and technical language?

The following paper firstly outlines the central methodological foundations. It then sets out the main findings of the project. The design criteria for language-sensitive geography teaching will be displayed, and insights into the effectiveness of language-sensitive geography teaching are given.

2 Methodology

Based on fundamental considerations of our research design in the context of design-based research, the sampling of our research project, as well as aspects of the data collection, will be described below.

2.1 Research frame

Methodically we approached the research questions set out above with design-based research (Bakker, 2018; McKenney & Reeves, 2019). It means that the two issues of (1) what language-sensitive geography teaching looks like and (2) how effective it is, are intertwined. We firstly developed design criteria for language-sensitive geography lessons based on the current state of research in cooperation with practice experts. The design criteria were then applied in a teaching unit, which is the subject of our study. The study was conducted in two design cycles. The knowledge gained in cycle I led to design cycle II, also consisting of conception, implementation of the teaching unit, data collection, and evaluation. Finally, the aim was to develop transferable standards for language-sensitive geography teaching, an empirically tested teaching unit, and findings on its effectiveness and its requirements for success (Figure 1).

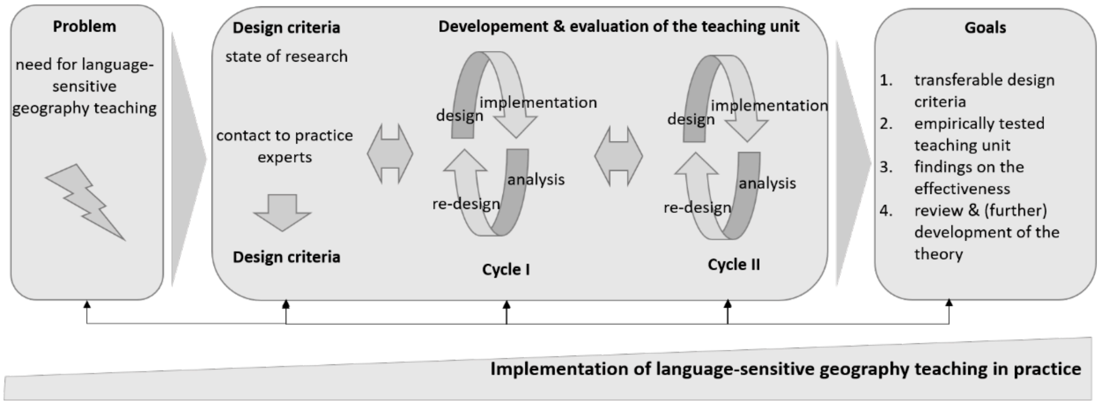

Figure 1: Overview of the research design

2.2 Sample

The required sample size was calculated with G*Power (Faul et al., 2007). A one-sided t-test for independent samples with the desired medium effect size of Cohen's d = .40, an α error of .05, and a power of .80 resulted in a sample size of 156 per design cycle. Thus a total of 320 students from twelve classes and six Berlin secondary schools took part in the study; the allocation of courses to experimental or control groups was randomized per school. To create a meaningful sample of the target population, namely 7^{th} graders, schools from different catchment areas in Berlin were selected. The example shows heterogeneity considering the first languages of the children, their socio-economic backgrounds, and their performances. The data presented in the following refer to the second design cycle, N = 185; the improved teaching unit and the enhanced survey instruments were used.

2.3 Data collection

The data collection was carried out in a pre-post-follow-up design. The data presented in the following focus on the pre-post comparison. We developed two teaching units, one experimental and one control group. Each received six lessons on the topic of the Earth's shell structure and plate tectonics but differed in the degree of language sensitivity. The experimental group received language-sensitive geography lessons, whereas control group I did not. Furthermore, a second control group was established to test whether learning gains could be achieved only by repeating the surveys (Figure 2). However, there were no significant learning gains in control group II; thus, it was left out in the second cycle.

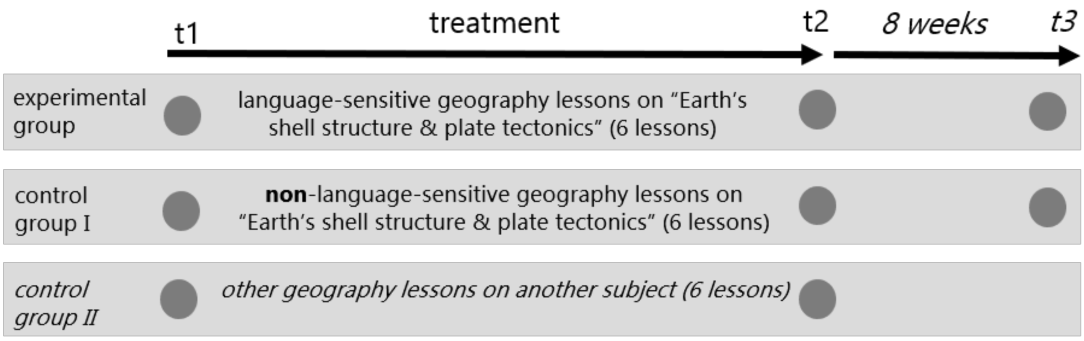

Figure 2: Research process per design cycle

The measurements focused on the two target variables knowledge and technical language, as well as the self-assessment of the students. Geographical knowledge on the subject of the Earth's shell structure and plate tectonics was assessed using a self-developed test (Cronbach's α = .79). The test aimed to be linguistically easily accessible. The target variable technical language was operationalized through two different survey instruments in order to address different facets of technical language. The C-test (Grotjahn et al., 2002), on the one hand, is a select type of gap test, which is primarily used in language acquisition research. We adapted the test for our research purpose using geographical texts about the above topic. The reliability turned out to be very good (Cronbach's α = .96). In the profile analysis (Grießhaber, 2016), on the other hand, the participants were asked to write a text about the formation of the Alps, given a sequence of images. The texts were then evaluated with regard to various surface characteristics (Cronbach's α = .82). Moreover, the self-assessment of the students' geographical knowledge and technical language competence was assessed on a five-level Likert scale. The items are based on large-scale school achievement studies (Cronbach's α = .92).

3 Findings

In the following, the initial results of the research project are set out. On the one hand, we give insight into the developed design criteria. On the other hand, we focus on the first results concerning the effectiveness of language-sensitive geography lessons.

3.1 Design-criteria

The design criteria were derived from the current state of research in a variety of neighboring research areas. We do not claim that the criteria are complete and they can be prioritized differently depending on the student's needs. Nevertheless, the following criteria turned out to be useful and good to operationalize in our research project. The effectiveness of the criteria was tested via the language-sensitive geography teaching unit (see Chapter 3.2).

1. Good geography teaching as a basis

The main goal of language-sensitive geography lessons is to give all pupils the best possible access to geography. Thus, the basis of language-sensitive geography teaching needs to be "good" geography teaching. Although there is not only one way to teach geography, there is a broad consensus that good geography teaching is based on a moderate constructivist understanding of learning (Rempfler, 2018a). Moreover, proper geography lessons include incorporating different geographical methods and the basic concepts of geography (Fögele, 2016; Rempfler, 2018b). The treatment adheres to the essential criterion of "good" geography lessons.

2. Scaffolding

Various empirical studies in educational research and second language didactics suggest that scaffolding has high effectiveness for language and subject learning in the classroom (Barricelli, 2015; Götze, 2018; Prediger & Zindel, 2017). Scaffolding is based on Vygotskij's "zone of proximal development" (Cole & Vygotskij, 1979). According to Vygotskij, learning is best stimulated when the aim is slightly above the current performance level. In order to reach the next level, scaffolds are offered and then gradually withdrawn. The treatment contains scaffolds on a linguistic level, for example, to relieve reading and writing on the one hand and to stimulate the corresponding learning process on the other.

3. Networking of representations

Especially in didactics of mathematics and in second language didactics, the networking of representations is considered to be fruitful for subject and language learning (Beese et al., 2014; Gibbons & Cummins, 2002; Wessel, 2015). The aim is to use different forms of representation to present the same content. The networking of the forms is one key to understanding. The language-sensitive teaching unit includes various forms of representations and offers to connect them.

4. Inclusion of the first language

Second language didactics point out the necessity of including the first language in subject learning. Although teachers cannot manage every language in their classroom, it is possible to consider the first language of the students as a potentially useful resource (Gogolin, 1994; Marx, 2014; Riegger et al., 2017). The students participating in the study were invited to use their first language if they felt it would be helpful for them.

3.2 Treatment effects

The teaching unit based on the criteria was examined, testing a set of specific hypotheses. Next, we outline our three main hypotheses and our evaluation strategy.

Following the logic that knowledge gain is only possible if one understands the information provided linguistically, it seems plausible that the experimental group (EG) should have higher increases in knowledge than the control group (CG). Since technical language is explicitly addressed in the experimental group, but not in the control group, we also expected advantages for the experimental group in the gain of technical language. Furthermore, we assumed changes in self-assessment in the experimental group since the self-assessment of the participants is an explicit part of the language-sensitive treatment (Table 1).

Table 1: Statistical hypotheses for pre-post-difference **Δ**

$\text{Effects}H_1\text{KnowledgeTest}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) > (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$	$\text{Effects}H_0\text{KnowledgeTest}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) \leq (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$
$\text{Effects}H_1\text{TechnicalLanguage}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) > (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$	$\text{Effects}H_0\text{TechnicalLanguage}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) \leq (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$
$\text{Effects}H_1\text{Self-assessment}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) \neq (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$	$\text{Effects}H_0\text{Self-assessment}$ $(\text{mean}(EGt_2) - \text{mean}(EGt_1))/((s(EG_{t2}) + s(EG_{t1}) / 2)) = (\text{mean}(CGt_2) - \text{mean}(CGt_1))/((s(CG_{t2}) + s(CG_{t1}) / 2))$

Notes:
H_1: statistical alternative
H_0: null hypotheses
m: mean
s: standard deviation
t_1: pre-measurement
t_2: post-measurement

To test the hypotheses, we defined a change score Δ as Cohen's d for independent samples subtracting the change in the experimental group minus the change in the control group. Therefore, a positive Δ indicates higher relative gains (or smaller losses) in support of the treatment.

Group differences Δ for the pre-post-differences were thus compared by Cohen's d_{emp} and an associated critical d-value (d_{crit}). We chose a power (β-1) of ≥ 80 % and a significance level of $\alpha < .05$. If $d_{emp} \geq d_{crit}$, the statistical alternative hypothesis is accepted with 80% certainty, and the null hypothesis is rejected with 95% certainty.

Regarding higher learning gains, we find significant advantages for knowledge and technical language in the experimental group. These findings support the alternative hypotheses. For the self-assessment, no significant group differences were found; the alternative hypothesis is rejected (Table 2). Δd_{crit} was computed with G*Power (Faul et al., 2007) in a sensitivity analysis for independent groups t-tests (one-sided for every test score except for self-assessment, which was two-sided). This evaluation strategy of planned contrasts is more efficient and specific than other statistical procedures and contributes this way to the statistical validity (Hager, 2004).

Table 2: Treatment effects for four test scores Δ t2-t1 ($\alpha = .05$, $\beta\text{-}1 = .80$)

test	Δd_{crit}	Δd_{emp}	$\Delta d_{emp} \geq \Delta d_{crit}$	EG/CG	N	ΔM t2-t1	SD	t_{crit}	df
knowledge	0.40	0.51	yes	EG	88	0.12	0.11	1.66	156
				CG	69	0.06	0.11		
C-test	0.40	0.51	yes	EG	88	0.34	0.31	1.66	156
				CG	69	0.19	0.30		
profile analysis	0.40	0.51	yes	EG	88	6.24	10.66	1.66	156
				CG	70	0.27	12.60		
self-assessment	.45	-0.28	no	EG	88	-0.12	0.45	1.98	155
				CG	69	0.00	0.40		

4 Conclusion

On the one hand, our research aimed to look at what design criteria language-sensitive geography lessons could have. The four principles we adopted from the current state of research and worked with were useful in our project. On the other hand, we demonstrated the effectiveness of language-sensitive geography teaching in a specific age group for a specific geographical subdomain. In this setting, the experimental group shows clear advantages over the control group in terms of both knowledge and technical language. In contrast, no significant changes could be found for the self-assessment. In other words: all students know more after the lessons than before and, students of the experimental group know more than their peers in the control group.

For both research questions, the design-based research framework can be an advantageous approach for other age groups and geographical subdomains, primarily through the iteration and the formative evaluation. Also, the close cooperation with teachers throughout the research process was fruitful. The survey instruments we developed and adapted are suitable in their validity and reliability for measuring the target variables. Throughout the design cycles, the instruments were optimized so that they can be used for further research.

4.1 Limitations

One limitation is the sample size. We chose a significance level of .05 and a power of .80 being able to detect effects higher than half a standard deviation but not smaller effects. Another aspect is that, although the design criteria worked in our setting, it remains questionable how exactly the criteria operate in other contexts, for example, on a different topic or different age groups. Design-based research examines entire designs and ultimately cannot filter out isolated factors that could explain the success or failure of an intervention. This paper focuses on the pre-post comparison of the data; calculations, including the follow-up-tests, are still to come.

4.2 Prospects

According to our research, language-sensitive geography teaching is key to giving students better access to knowledge and technical language. However, there are still many questions to answer. For instance, further research on the changes in self-assessment should be done, especially in conjunction with the development of knowledge and language skills. In the case of interpreting the findings considering the self-assessment, we face difficulties. For now, we cannot say that the self-assessment of one or both groups has improved or worsened. It can be said that the experimental group assesses its skills as lower than before the treatment. One interpretation could be that the participants overestimated their abilities in the first assessment. After the treatment, which focuses on linguistic difficulties, the self-assessment could have become more realistic. It could also mean that the participants' self-assessment decreased because they felt overwhelmed by the treatment. In the analyses presented here, the entire sample of cycle II was considered. It might also be useful to examine various reference groups, separated by gender and reading literacy, to investigate differential effects depending on the group. Moreover, calculations on the stability of the effects including the third measurement time are required.

Acknowledgments

We would like to thank Dr. Barbara Geist for her valuable advice on this research project. We thank the Senatsverwaltung of Berlin for approving our research, the involved headmasters, teachers, pupils, and their parents who gave consent for them to participate and our colleagues from the Chair of Geography Education of FAU. The latter gave numerous hints for this study in fruitful team meetings. Many thanks also go to the Heinrich Böll Foundation that founded this project. Finally, we would like to thank Caroline Tremble for proofreading this paper.

References

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge.

Barricelli, M. (2015). Worte zur Zeit. Historische Sprache und narrative Sinnbildung im Geschichtsunterricht. *Zeitschrift Für Geschichtsdidaktik*, *14*, 25–46.

Beese, M., Benholz, C., Chlosta, C., Gürsoy, E., Hinrichs, B., Niederhaus, C., & Oleschko, S. (2014). *Sprachbildung in allen Fächern*. *Deutsch lehren lernen: Vol. 16*. Klett-Langenscheidt.

Cole, M., & Vygotskij, L. S. (Eds.). (1979). *Mind in society: The development of higher psychological processes* (2. printing). Harvard Univ. Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fögele, J. (2016). *Entwicklung basiskonzeptionellen Verständnisses in geographischen Lehrerfortbildungen*. *Geographiedidaktische Forschungen: Vol. 61.*

Gibbons, P., & Cummins, J. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Heinemann.

Gogolin, I. (1994). *Der monolinguale Habitus der multilingualen Schule*. Waxmann.

Götze, D. (2018). Schriftliches Erklären operativer Muster fördern. *Journal Für Mathematik-Didaktik*, *35*(1), 147. https://doi.org/10.1007/s13138-018-00138-4

Grießhaber, W. (2016). *Diagnostik und Förderung leicht gemacht*. *Deutsch als Zweitsprache in der Grundschule*. Ernst Klett Sprachen.

Grotjahn, R., Klein-Braley, C., & Raatz. U. (2002). C-Test: an overview. In J. Coleman, R. Grotjahn, & Raatz. U. (Eds.), *University language testing and the C-Test* (pp. 93–114).

Hager, W. (2004). *Testplanung zur statistischen Prüfung psychologischer Hypothesen: Die Ableitung von Vorhersagen und die Kontrolle der Determinanten des statistischen Tests*. Hogrefe Verl. für Psychologie.

Handro, S. (2018). Sprachbildung im Geschichtsunterricht: Leerformel oder Lernchance. In K. Grannemann, S. Oleschko, & C. Kuchler (Eds.), *Sprachbildung im Geschichtsunterricht: Zur Bedeutung der kognitiven Funktion von Sprache* (pp. 13–42). Waxmann.

Härtig, H. (2010). *Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests*. *Studien zum Physik- und Chemielernen: Vol. 101*. Logos Verlag.

Höttecke, D., Ehmke, T., Krieger, C., & Kulik, M. A. (2017). Vergleichende Messung fachsprachlicher Fähigkeiten in den Domänen Physik und Sport. *Zeitschrift Für Didaktik Der Naturwissenschaften*, 1–17. https://doi.org/10.1007/s40573-017-0055-6

Marx, N. (2014). Häppchen oder Hauptgericht? Zeichen der Stagnation in der deutschen Mehrsprachigkeitsdidaktik. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, *19*(1).

McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research* (Second edition).

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (Eds.). (2016a). *Pirls 2016: International Results in Reading.*

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (Eds.). (2016b). *TIMSS 2015: International Results in Mathematics.*

OECD (Ed.). (2019). *PISA 2018 Results (Volume I, II & III): Combined executive summaries.*

Paetsch, J., Felbrich, A., & Stanat, P. (2015). Der Zusammenhang von sprachlichen und mathematischen Kompetenzen bei Kindern mit Deutsch als Zweitsprache. *Zeitschrift Für Pädagogische Psychologie*, *29*(1), 19–29.

Prediger, S., & Hein, K. (2017). Learning to meet language demands in multi-step mathematical argumentations: Design research on a subject-specific genre. *European Journal of Applied Linguistics*, *5*(2), 1–27. https://doi.org/10.1515/eujal-2017-0010

Prediger, S., & Zindel, C. (2017). School Academic Language Demands for Understanding Functional Relationships: A Design Research Project on the Role of Language in Reading and Learning. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*, 4157–4188. https://doi.org/10.12973/eurasia.2017.00804a

Rempfler, A. (2018a). Wirksamer Geographieunterricht - Eine Einführung. In A. Rempfler (Ed.), *Unterrichtsqualität: Band 5. Wirksamer Geographieunterricht* (pp. 19–24). Schneider Verlag Hohengehren GmbH.

Rempfler, A. (2018b). Wirksamer Geographieunterricht - Eine Synthese aus 18 Experteninterviews. In A. Rempfler (Ed.), *Unterrichtsqualität: Band 5. Wirksamer Geographieunterricht* (pp. 206–219). Schneider Verlag Hohengehren GmbH.

Riegger, M., Bülow, A., Wlossek, I., & Rost-Roth, M. (2017). Berücksichtigung von Herkunftssprachen im schulischen Kontext: Erfahrungen und Einschätzungen von Schülerinnen und Schülern. In I. Fuchs, S. Jeuk, & W. Knapp (Eds.), *Beiträge aus dem Workshop "Kinder mit Migrationshintergrund". Mehrsprachigkeit: Spracherwerb, Unterrichtsprozesse, Seiteneinstieg: Beiträge aus dem 11. Workshop "Kinder mit Migrationshintergrund", 2015* (pp. 155–171). Fillibach bei Klett.

Scheuer, R., Kleffken, B., & Ahlborn-Gockel, S. (2010). Experimentieren als neuer Weg zur Sprachförderung. In H. Köster (Ed.), *Handbuch Experimentieren* (2nd ed., pp. 91–114). Schneider Hohengehren.

Schmiemann, P. (2011). Fachsprache in biologischen Testaufgaben. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *17*, 115–136.

Schmölzer-Eibinger, S. (2013). Sprache als Medium des Lernens im Fach. In M. Becker-Mrotzek (Ed.), *Fachdidaktische Forschungen: Bd. 3. Sprache im Fach: Sprachlichkeit und fachliches Lernen* (pp. 25–40). Waxmann.

Wessel, L. (2015). *Fach- und sprachintegrierte Förderung durch Darstellungsvernetzung und Scaffolding: Ein Entwicklungsforschungsprojekt zum Anteilbegriff. Dortmunder Beiträge zur Entwicklung und Erforschung des Mathematikunterrichts: Vol. 19*. Springer Spektrum.

Contact email: santina.wey@fau.de