ARE TEACHERS' LANGUAGE VIEWS CONNECTED TO THEIR DIAGNOSTIC JUDGMENTS ON STUDENTS' EXPLANATIONS?

Susanne Prediger, Dilan Şahin-Gür & Carina Zindel
TU Dortmund University, Germany

When teachers analyze students' explanations in a language-responsive mathematics classroom, they explicitly and implicitly activate various categories and hold different views on language and mathematics learning. This study investigates how typical views on language in mathematics classrooms are related to what teachers consider relevant in their diagnostic judgments, in this case on students' explanations of the slope formula for linear functions. Seventy-eight teachers' personal constructs were elicited using a diagnostic activity and related to their self-reported views on language and mathematics learning. The group of language reducers can be shown to focus significantly more on the surface level of language whereas the language pushers *group focuses more on the discourse level. In contrast, worries about language responsiveness being time consuming do not seem to influence diagnostic judgments.*

Due to the increasing language diversity in mathematics classrooms, fostering language learners' mathematics learning has become a major task for mathematical teachers all over the world. Although the teachers' crucial role in language-responsive classrooms has often been acknowledged in classroom research, research surveys show that few studies have been conducted on *teachers'* resources and the obstacles they face in developing language-responsive mathematics classrooms (Radford & Barwell, 2017; Barwell et al., 2016)

In this paper, mathematics teachers' resources and their obstacles when developing their language-responsive classrooms are conceptualized by typical *views* (on language and mathematics learning) and mathematics- and language-related *categories* they apply when noticing students' products (following Sherin, Jacobs, & Philipps, 2001). We report on an inquiry with mathematics teachers (n = 78) that investigates their personal categories elicited in comparative diagnostic judgments which pursues the following research questions:

Which categories do teachers activate when conducting diagnostic judgments on students' written explanations of slope?

How are these categories connected to the teachers' views on language?

Prediger, Susanne, Şahin-Gür, Dilan & Zindel, Carina (2018). Are Teachers' language views con-nected to their diagnostic judgments on students' explanations? In Ewa Bergqvist, Magnus Österholm, Carina Granberg, and Lovisa Sumpter (Eds.). Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education (Vol. 4, pp. 11-18). Umeå, Sweden: PME.

THEORETICAL BACKGROUND: TEACHERS' PERSONAL VIEWS AND CATEGORIES IN LANGUAGE-RESPONSIVE CLASSROOMS

As Sherin et al. (2001) have argued, teachers' classroom practices rely heavily on what they notice in classroom complexity. Noticing in language-responsive mathematics classrooms starts with diagnostic judgments on students' utterances and products, especially on students' written explanations of mathematical concepts, because explaining mathematics concepts is the most important discourse practice, especially for language learners (Moschkovich, 2010).

Personal views and personal categories in teachers' diagnostic judgments

Sherin et al. (2011) describe noticing as a complex process "through which teachers manage the blooming, buzzing confusion of sensory data with which they are faced" (p. 5). Hence, noticing selected aspects involves filters of perception and consists of two processes: "attending to particular events in an instructional setting" and "making sense of events in an instructional setting" (ibid., p. 5). For the second process of making sense, they hint at the relevance of categories: Interpreting means "relating observed events to abstract categories" and characterizing what they see in terms of familiar phenomena (ibid., p. 5). As Prediger and Zindel (2017) have shown that extrapolating the personal categories in teachers' diagnostic judgments can be an interesting research approach for unpacking their thinking and noticing. As the personal categories that teachers activate in diagnostic judgments can be widespread and very individual, they cannot be captured by predefined items but should be elicited with open-ended diagnostic activities.

Beyond the specific categories, teachers' diagnostic judgments can be influenced by their general views on the topic in view, in our case views on the role of language in mathematics classrooms and the individual interpretation of what language-responsiveness might mean (Short, 2017).

Categories and language views relevant in language-responsive classrooms

Although Moschkovich (2010, p. 160) recommended investigation of teachers' judgements in language-responsive classrooms, little so far is known on teachers' personal views and categories about language and mathematics learning. This can be partly explained by the fact that most research studies on language are conducted as classroom observation studies (Radford & Barwell, 2016).

Although teachers' diagnostic judgments have not been investigated directly, these classroom observations provide important hints on potentially crucial categories and views. As the research on language in classroom observation studies has shown, the role of language in mathematics classrooms cannot be adequately grasped on the *word level* alone, as it is tightly connected to the *sentence level* and especially the *discourse level*: Classroom studies in linguistics and mathematics education research have shown that the epistemic function of language (i.e., the role of language for higher order thinking practices) is mainly reflected by participation in classroom *discourse practi-*

ces: Language learners in their early stages of language proficiency can have difficulties participating in classroom practices such as explaining meanings or describing general patterns, whereas they can participate in reporting procedures or describing by examples (Moschkovich, 2010). Thus, categories on the discourse level are crucial for focusing relevant phenomena, as they allow for integrating language and mathematics learning in much deeper way than simple categories on the language surface level (such as orthography) or the word level (such as identifying relevant technical terms).

A typical view that might hinder language learning concerns the interpretation of language-responsiveness as language simplification: In order not to exclude language learners from mathematics learning, many teachers tend to reduce the language demands by simplifying all texts and reducing the production expectations to keywords and half sentences. In contrast, language education research has emphasized the requirement of comprehensible but demanding language input and pushing much language output in order to enhance language learners' learning opportunities (e.g., Short, 2017). Pushing language in the zone of proximal development thus seems to be an important overall view on language in classrooms.

In general, each classroom innovation can be hindered by teachers' view that this innovation is too time consuming. Understanding the backgrounds of time worries might therefore help to overcome them and increase the chance that teachers adopt approaches of language responsiveness. Existing case studies have led to the hypothesis that fewer time worries may be held by teachers who see how language and mathematics are deeply connected, which means that they are already addressing the discourse level on which content- and language-integration mainly occurs.

In order to investigate how categories and views are connected, the following two hypotheses are tested in this study:

H1 Those teachers who worry that language responsiveness is time consuming focus less on the discourse level and more on surface levels than those who do not worry.

H2 Those teachers who try to reduce language in their classroom focus less on the discourse level and more on the surface levels than those who push language.

METHODS

Methods of data gathering

Sample. The sample consisted of German middle and high school mathematics teachers ($n = 78$) in their first session of a volunteer professional development series on language-responsive mathematics classrooms. The teachers had 2-30 years of experience in math teaching (with a median of 6-10 years) and between 0 hours and several days of previous encounters with ideas of language-responsive classrooms (with a median of 6-8 hours).

Questionnaire for general views on language in math classrooms. General views were captured using e.g. the following items:

- *Language reducer vs. language pusher:* "For language learners, I try to reduce the language."
- *Time worrier vs non-worrier:* "Language responsiveness is an additional task for math classrooms which steal us much time from mathematics learning."

Teachers' views captured on the six-point Likert-type scales allowed the formation of sub-samples of *language reducer* vs. *pusher* and *time worrier* vs. *non-worrier*, containing those teachers who selected strongly disagree/quite disagree or quite agree/ strongly agree, without those who chose partially dis-/agree.

Diagnostic activity. Eliciting the teachers' implicit personal categories on language and mathematics followed the variation principle: Diagnostic judgments were requested for three contrasting students' explanations (Fig. 1). Teachers were asked to name their criteria, evaluate the three texts according to them, and justify their evaluation.

Diagnostic Activity

The Grade 8 class has introduced the meaning and the formula for the slope of linear functions.
The homework task (formulated as on the right) requested to write a summary.

Explain what $m = \frac{y_2-y_1}{x_2-x_1}$ means and what you can describe with it.

Analyze the three students' texts using four self-defined categories, two for mathematical aspects and two for language aspects. Evaluate by 0,1 or 2 points and justify your decisions.

Ali	Suleika	Tom
You want to calculate the slope. First you chose two points, e.g. x=3, y=1 and x=5, y=8. Second, you evaluate: $m = \frac{8-1}{7-3} = \frac{7}{2}$. Ready	The slope saying, how much growing the function per x-step. Thus, how much get y more per how much get x more. The PER makes the DIVIDED. 	Iff yu for exempel have y=0.2+10 of the mobile phone tarif. For exempel at 10 minutes it is 12 €, at 30 minutes it is 16 €. Thus, distance 20, price more 4 € and than $\frac{4\,€}{20\,min} = \frac{0.2\,€}{1\,min}$. Thenn, costs per minute my price 0.20 € more costly.

Fig. 1. Diagnostic activity for teachers (translated from German with errors preserved)

The slope of a linear function provides a mathematically rich exemplary topic for students' explanations. This topic demands not only procedural knowledge when evaluating the slope formula for specific values but also conceptual knowledge explaining its meaning as a whole (the slope captures how much a function grows) as well as the components of the quotient: The ratio of two distances is an interpretation that requires conceptual understanding of different arithmetic models (Usiskin, 2008). Four discourse practices can be distinguished here, reporting procedures, explaining meanings, general phrasings, and concrete phrasings (Suleika explains meanings in a general way, whereas Ali reports procedures concretely). The three students' texts were chosen to show a wide spectrum of language features on the surface level (e.g., orthography), word level (technical terms), sentence level (grammatical structures), and discourse level (with the four discourse practices mentioned).

Data analysis procedures

The manifold personal criteria that teachers stated for the diagnostic activity in Fig. 1 were analyzed by a specifically developed categorial scheme. The first version of the categorial scheme was derived from the current state of research and then adapted to the data in order to capture all personal criteria. As the teachers used the same words for criteria with different individual meanings, the verbatim criteria, their assessment scores for each student text and their justifications also had to be taken into account for the categorization. Table 1 shows examples for the categorizations, and Table 2 shows the complete categorial scheme. Within the categorized data, frequencies of category use were determined for the whole sample and compared for the sub-samples. In order to test Hypotheses H1 and H2 in terms of the differences of the sub-samples, *t*-tests were administered.

RESULTS

Insights into two cases

The cases in Table 1 show that the diagnostic activity can elicit very different personal criteria, as intended: The two teachers (here called Peter Tremnitz and Anne Schäfers) assess the students' explanations differently (the bold numbers indicate the evaluations they assigned to Ali, Suleika, and Tom), these assessments scores are based on different personal constructs underlying their diagnostic judgments. Some teachers' criteria are categorized under more than one category, mostly because their justification address several aspects. These personal criteria vary between very vague aspects such as mode of expression and core categories on the discourse level. It is typical that criteria on the discourse level (general/concrete phrasing, explaining meanings/reporting procedures) appear sometimes as mathematical criteria, sometimes as language criteria.

	Mathematical Criterion A	**Mathematical Criterion B**	**Language Criterion C**	**Language Criterion D**
Peter Tremnitz language pusher time non-worrier	**202** "Which meaning (calculating slope): finding 2 points and using formula" *(→ procedural knowledge)*	**111** "What to describe (difference quotient)": Word slope. meaning of quotient and using (the word) slope / difference quotient is developed for an example" *(→ conceptual knowledge, technical terms, concrete phrasing explaining meaning)*	**021** "Explanation of slope (Why-question): does not explain, only describe, row-per-step / more per…" *(→ conceptual knowledge, explaining meaning, technical terms)*	**101** "How to use formula? (How-question): abstraction is missing" *(→ procedural knowledge, describing procedure, general phrasing)*
Anne Schäfers language reducer partial time worrier	**121** "mathematical correctness of formulations" *(→ mathematical correctness)*	**222** "application" *(→ procedural knowledge)*	**110** „technical language: slope, difference" *(→ technical terms, surface level)*	**210** "understandability" *(→ understandability)*

Table 1. Examples of evaluations and elicited criteria with justification of two teachers (**in boldfaced type**: assessment scores for Ali, Suleika, Tom*; in italics*: categories assigned by researcher)

Peter has a strong focus on the connection between language and mathematics in three of his four criteria. When he mentions technical terms, they serve as indicators for deeper aspects on the discourse level, as his main distinctions concern the different discourse practices. In contrast, Anne presents disconnected criteria without addressing the discourse level at all. Interestingly, these findings correspond to their views expressed in the self-report scale in the hypothesized ways: Peter favors views as language pusher and does not really worry about time, whereas Anne is a language reducer and tends to partially worry about language responsiveness being time consuming.

In the accompanying group discussion, the connection became also apparent: As language for her is only located on the surface level and not really connected to mathematics, it is rational in her view to reduce language demands. In contrast, for Peter, the mathematics and language criteria are tightly connected, so language is to be pushed to foster mathematics learning. Also for other aspects, the insights provided by the questionnaire resonate with richer qualitative video data from group discussions with these two teachers.

However, even if these two cases resonate with the hypotheses, the hypotheses must be tested for a larger sample.

Quantitative results of typical categories and connections

Table 2 presents the frequencies of categories built from the elicited personal criteria as exemplified in Table 1.

Teachers' categories with different focus	**Frequency of category use… …in whole sample**	**…among language reducers – pushers**	**Effect size *d***	**… among time worriers – non-worriers**	**Effect size *d***
Focus on mathematical criteria					
Mathematical correctness	24%	**53% – 06%****	**1.18**	09% – 14%	0.16
Conceptual knowledge	73%	65% – 72%	0.16	91% – 77%	0.37
Procedural knowledge	54%	**35% – 78%****	**0.92**	55% – 49%	0.12
Focus on discourse level (Language or Math)	*39%*	38% – 60%	0.43	40% – 42%	0.05
Concrete, example-bound phrasing	26%	80% – 57%	0.46	75% – 67%	0.17
General phrasing	26%	40% – 71%	**0.61**	50% – 67%	0.31
Explaining meaning	32%	**12% – 44%***	**0.76**	18% – 37%	0.42
Reporting procedure	16%	18% – 22%	0.11	18% – 17%	0.03
Focus on Language beneath discourse level					
Only surface level or very vague criteria	65%	**80% – 53%***	**0.71**	70% – 70%	0.01
Orthography	12%	13% – 06%	0.24	18% – 18%	0.01
Mode of expression	20%	06% – 22%	0.46	18% – 26%	0.19
Understandability	26%	31% – 17%	0.34	09% – 26%	0.46
Word level: technical terms	42%	50% – 33%	0.33	45% – 47%	0.03
Sentence level: syntactical issues	23%	31% – 17%	0.34	36% – 26%	0.21

Differences between subsamples that are significant in *t*-test are marked in bold with * for $p < .05$ and ** for $p < .01$. Medium effect sizes ($d > 0.5$) and high effect sizes ($d > 0.8$) are marked in bold even if not significant.

Table 2. Frequencies of different categories: Comparison of sub-samples

In the whole sample, 65% of the teachers address only very vague or surface criteria for language, and 42% adopt the often criticized focus on isolated technical terms. Interestingly, 39% of the teachers address categories on the discourse level, distinguishing in some ways between general and concrete phrasing and/or reporting procedures and explaining meanings. Interestingly, half of these teachers mention discourse practices as a mathematical instead of a language criterion.

The comparison of sub-samples shows that the pattern exemplified by the two cases only partly re-appear: For the sub-samples of time worriers and non-worriers, the frequencies of categories are similar without any significant differences (with a maximal difference of 19% for explaining meanings) and all have small effect sizes ($d < 0.47$ for all categories). Thus, Hypothesis H1 must be rejected: Worries about language responsiveness being time consuming does not seem to be systematically connected to the personal constructs applied for diagnostic judgments.

In contrast, the language reducers and language pushers have significantly different priorities in their diagnostic judgments: Whereas the language reducers often address the very general category of mathematical correctness, the language pushers differentiate more thoroughly between procedural and conceptual knowledge. 80% of the language reducers focus exclusively on surface levels (including orthography or technical terms), whereas only 53% of the language pushers do. In contrast, 60% of the language pushers focus on the discourse level while only 38% of the language reducers do; the difference is specifically significant for the most important discourse practice of explaining meanings. So, Hypothesis H2 can be confirmed, with the corresponding null hypothesis being rejected.

DISCUSSION

This study followed Moschkovich's (2010, p. 160) recommendation to investigate teachers' judgements in language-responsive classrooms, in this study, the diagnostic judgments on explanations of a mathematical concept. Similar to other investigations of teachers' diagnostic judgments (Prediger & Zindel, 2017), the thorough exploration of individual categories turned out to provide insightful windows into teachers' thinking.

The empirical identification of teachers' personal categories revealed the problem of surface level categories being dominant for 65% of the teachers. For these teachers, the recommended shift of focus to the discourse level should be a crucial part of professional development programs (Moschkovich, 2010; Short, 2017). This is specifically important as the general view of language responsiveness as language reduction turns out to be significantly connected to the missing focus on the discourse level. As long as professional development programs fail to address the discourse level (as criticized by Moschkovich, 2010), this study provides indications that a crucial precondition for an adequate view on language-responsiveness is missing.

While 39% of the teachers in the study already paid some attention to the discourse level, half of these teachers mention discourse practices as a mathematical instead of a language criterion. So this focus of attention reveals an important resource on which professional development programs must build upon in order to shift teachers' language focus to the discourse level.

Teachers' attitudes have often been described as quite stable orientations that have an impact on teachers' noticing and practices. In this paper, the relation may be turned around: For a new challenge, such as language-responsive mathematics classrooms, the presented findings provide indications that the scope of teachers' categories may also influence their personal views on the role of language in mathematics classrooms. However, this seems to apply more for language reducers (Hypothesis H2 was confirmed) than for those who worry about language responsiveness being time consuming (Hypothesis H1 had to be rejected).

In order to overcome limitations of the study, future research will increase the (so far limited) sample size and study also the development of teachers' views and personal categories during a PD program, with the survey and also qualitative means.

Acknowledgment. The study was conducted with the DZLM-project MuM-implementation (BMBF grant 03VP02270 to S. Prediger by the German Ministry for Education and Research). We thank Susanne Willer and all student researchers for assistance in the analysis and Henrike Weinert for the statistical support she conducted.

References

Barwell, R., Clarkson, P., Halai, A., Kazima, M., Moschkovich, J., … Villavicencio, M. (2016). *Mathematics education and language diversity: The 21st ICMI Study*. Dordrecht: Springer.

Moschkovich, J. N. (2010). Recommendations for Research on Language and Math. Ed. In J. Moschkovich (Ed.), *Language and Mathematics education* (pp. 1-28). Charlotte: Infor. Age.

Prediger, S., & Zindel, C. (2017). Deepening prospective mathematics teachers' diagnostic judgments. *European Journal of Science & Mathematics Education, 5*(3), 222-242.

Radford, L., & Barwell, R. (2016). Language in mathematics education research. In A. Gutiérrez, G. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 275-313). Rotterdam: Sense.

Usiskin, Z. (2008). Arithmetic Curriculum and the Real World. In D. Bock, B. Søndergaard, B. Gómez, & C. Cheng (Eds.), *Proceedings of ICME11, TSG10* (pp. 9-16). Monterrey: ICMI.

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011). *Mathematics Teacher Noticing*. New York: Routledge.

Short, D. J. (2017). How to Integrate Content and Language Learning Effectively for English Language Learners. *EURASIA Jour. of Math., Science & Techn. Ed*. 13(7b), 4237-4260.