

5. Discussion

This investigation explored the base literacy skills required for word reading and text comprehension among bilingual and monolingual early readers. The longitudinal data set allowed for the investigation of mean differences in base skills, deviations in predictive models, differential growth patterns, and the analysis of model fit for both groups. Overall, the most significant group differences emerged in relation to the expressive vocabulary skills scales, in which children in the bilingual group identified substantially fewer German words. No meaningful differences were found in reading performance between the two groups. However, the predictive factors varied in their power to explain the children's levels of reading comprehension. The proposed theoretical model of reading required considerable modification to fit both groups, resulting in unique models for each group. This chapter assesses the overall quality of the instruments and highlights the most important characteristics of the sample. Each section of the chapter outlines the essential findings of the study with a discussion of their implications for the four primary hypotheses. Finally, this chapter provides suggestions for further research and discusses the practical implications of this investigation for educational programs and policies geared toward minority language children in Germany.

5.1. Preliminary analyses and descriptives

5.1.1. Instrument characteristics

Because the majority of the measurement scales in this investigation were utilized for the first time in either a German or a Turkish-German setting, the psychometric properties of the central instruments were assessed for both the bilingual and monolingual groups. Overall, the measures were determined to be acceptable for the current research purposes. This section evaluates only the most critical findings of the instrument analyses.

The revised *phonological awareness scales*, altered to be linguistically neutral for native speakers of either Turkish or German, appeared to function well with both samples. Scales correlated significantly with each other at each point in time and were relatively stable over six months. In confirming the unidimensionality of the instrument, it was found that the four phonological awareness sub-scales factored onto a single phonological awareness scale for both groups. The phonological awareness measure demonstrated satisfactory factor analytical structures and high reliability.

The revised *pseudoword memory scale* did not demonstrate unidimensionality. Two factors emerged for both the bilingual and monolingual groups. A larger factor consisted of shorter/simpler items and a smaller factor contained the longer/difficult items. Although there

were some slight inconsistencies in individual factor loadings, the instrument was deemed acceptable for use in this investigation due to the good reliability of the instrument and the easily interpretable two-factor structure.

The *German vocabulary measure*, adapted from a U.S.-developed bilingual verbal abilities instrument, proved to function well in a German-speaking environment. The two longitudinally administered scales were very stable over the six months between testing sessions. A factor analysis showed that for both groups, a single clear factor accounted for around 50% of the variance in each subscale. To test the external validity of the vocabulary measures, correlations with teacher assessments of German language abilities were conducted and found to be generally strong and similar across groups. The instrument demonstrated equally strong psychometric properties for both German monolingual and Turkish bilingual children.

A series of bivariate correlations, conducted to better understand the linguistic development of the bilingual and monolingual samples, illustrated the associations between home and educational experiences and language abilities in both Turkish and German. Interestingly, for the German monolingual group, the analyses indicated no significant correlations between the vocabulary scales and home or educational variables such as parental reading practices or preschool attendance. Based on the bivariate correlation analyses, neither group demonstrated associations between German vocabulary skills and parents' reading aloud practices. Turkish vocabulary skills appeared to be negatively associated with attending after-school daycare and, not surprisingly, positively correlated with the reported amount of Turkish spoken in the home. Overall, it was unexpected that extra-curricular variables would play such a minimal role in vocabulary performance in the second grade for all children in the sample. A viable explanation for the lack of home influence on vocabulary scores could be that the vocabulary measures were based on cognitive academic language proficiency (CALP) which is, by definition, used primarily in academic or school environments (Munoz-Sandoval et al., 1998).

The revised Knuspel *listening comprehension test* (Marx, 1998) posed several difficulties. Although it is a well-known standardized test for children in Germany, shortened only by two items for this investigation, the listening comprehension scale demonstrated no clear factor analytic structure. Additionally, the two groups differed erratically on the factor loadings, producing no interpretable patterns. Furthermore, the assumption that listening comprehension is an integral part of verbal abilities was discounted by the low correlations with the vocabulary measures. It is not clear if the scale did not function due to the modified method used to score the instrument in this study or if the scale in itself is generally weak. Listening comprehension was therefore not used as a central measure of verbal abilities, but more as an exploratory factor in this investigation.

Several significant differences emerged in the scale intercorrelations between the bilingual and monolingual groups. Although compared to the vocabulary measures, the phonological awareness scales correlated more strongly with reading performance in both groups, the correlation between reading and phonological awareness was significantly stronger for the monolingual group at the end of second grade. This divergent correlation can be interpreted as a preliminary indication of possible differences between monolingual and bilingual children in the factors responsible for successful reading comprehension. The more in-depth analyses performed to explore that possibility will be discussed in section 5.3..

5.1.2. Participant language and learning patterns

To overcome a methodological weaknesses found in most studies involving minority language children and academic performance, a rigorous process was undertaken to ensure that the two comparison groups were as homogeneous as possible. One group was composed of monolingual German speakers, exposed only to German in the home (as reported by both teachers and students). The other consisted of bilingual Turkish-German speakers, exposed only to Turkish (and often German) in the home. Children who reported that languages other than Turkish and/or German were spoken in the home were eliminated from the sample. Unique for research of this nature, this investigation collected data on the L1 abilities of participants to check for proficiency in the two languages as an inclusion criterion for the bilingual group. Many studies classify all minority language students together in an “L2” learner category without examining the children’s actual language abilities in both languages and without regard for the critical differences between home languages and cultures. This procedure neglects valuable information about the verbal abilities and linguistic experiences of the minority language participants. By taking a closer look at these aspects with teacher assessments and individual participant measures, the present study provides a more comprehensive depiction of the linguistic abilities of the children under investigation and a more linguistically homogeneous group of bilingual children than is typically available.

Teacher assessments of the children’s German language skills, readiness to learn, and concentration abilities produced several interesting findings on the linguistic and academic characteristics of the participants. Turkish-German bilinguals were rated as having weaker German skills both in first and in second grades, but they were also rated by their teachers as possessing a significantly stronger readiness to learn. Similar findings are often reported for minority language students in U.S. studies (e.g., Suarez-Orzoco & Suarez-Orzoco, 1995), but have not been researched in depth among young Turkish-German students.

Conducting Turkish language tests parallel to German language tests allowed for a comparison of abilities in the two languages in the two groups. For the Turkish-German bilinguals, German was clearly the stronger language at both points of measurement in the second grade. Furthermore, the results suggest that while German abilities develop significantly through second grade, this is not the case for Turkish vocabulary, which shows more of a stagnation. Although data were available for two points of measurement only, these findings indicate that the instruction provided to children in the urban Berlin schools promotes German expressive vocabulary skills more strongly than the family environment promotes Turkish vocabulary skills.

As a further step in the description of vocabulary skills in the bilingual sample, the associations between scores for each scale in Turkish and German were examined. There were very few significant correlations between Turkish and German vocabulary skills. The ability to find synonyms and antonyms across the two languages correlated moderately. These findings do not provide any support for the theory of language interdependence, in that the vocabulary abilities in the two languages correlated only minimally (in contrast to Cummins, 1991).

Finally, the relationships between scores on the Turkish and German vocabulary measures and all other primary instruments were examined to investigate the possibility of cross-language transfer of language abilities in the L1 to reading skills in the L2 for the bilingual group. Two interesting findings emerged. First, although measures of German vocabulary correlated significantly with all measures of phonological awareness, decoding, and reading comprehension at all times of measurement, Turkish vocabulary skills correlated with none of the verbal or reading related scales. This indicates that there is no apparent cross-language transfer in the bilingual group of Turkish-German children from Turkish to German. It should be pointed out, however, that being a speaker of Turkish was not negatively related to performance in any of these areas either. Moderate correlations between phonological awareness and verbal abilities were expected, however, since the phonological test was created to be non-word based and language neutral, it was unexpected that phonological awareness was not equally related to German and Turkish vocabulary abilities. This may be due to the fact that, with the exception of Turkish verbal abilities, the measures were all administered in German. To examine this explanation, an additional group of children would need to be provided with the test instructions in Turkish. Although their measure of phonological awareness was an English test, Carlisle and collaborators (1999) reported similar findings for their Spanish-English bilingual sample: phonological awareness correlated highly with English vocabulary skills, but not at all with Spanish vocabulary skills.

The second finding in the analyses aimed at investigating how cross-language skills among the bilingual group were related to the correlations between vocabulary and cognitive abilities. Turkish vocabulary skills correlated with cognitive abilities (administered in German) in the middle of the second grade, but the relationship essentially disappeared by the end of second grade. German vocabulary, however, maintained its moderate correlation with cognitive abilities throughout the second grade. One interpretation of that finding could be that the non-verbal cognitive abilities test (the “Culture Free Test”) was more strongly associated with German abilities and perhaps not as “culture free” as its name suggests. Again, the administrators of the cognitive abilities test provided the test instructions in German. It is plausible that this led to a stronger relationship between cognitive abilities and German verbal abilities than if the test instructions had been non-verbal or provided in both Turkish and German.

In sum, although only vocabulary skills were measured across both languages, it does not seem that language transfer or interdependence is a substantial part of the bilingual Turkish-German children’s linguistic profile. Proctor and colleagues (2005) proposed that L1 abilities are perhaps the missing crucial information for the prediction of L2 reading processes of bilingual readers. This does not appear to be the case for the sample at hand. Based on the scale intercorrelations, it seems unlikely that the L1 vocabulary skills of the bilingual participants play a meaningful role in their German reading comprehension or component reading skills. Nonetheless, having a picture of the verbal abilities of the bilingual children in both languages is invaluable in understanding their linguistic development and needs.

5.2. Mean differences between bilingual and monolingual readers

Hypothesis 1 predicted that several significant mean differences would emerge when comparing the Turkish-German bilingual and German monolingual readers on their core reading abilities. The first sub-hypothesis regarding mean differences in performance presented in this thesis posited that bilingual children with command of both Turkish and German would have a heightened sense of phonological awareness (Hypothesis 1a). This hypothesis was drawn from a range of literature and theories pointing toward better metalinguistic skills among bilingual children as well as findings showing particularly advanced phonological skills among monolingual Turkish children. Therefore, it was predicted that a population with both of these attributes would show better phonological awareness as well. In line with this hypothesis, the bilingual group did, in fact, perform significantly better on one phonological subscale at the end of second grade, and scored consistently, though marginally, better than the monolingual group on the other subscales at each time of measurement. Although the bilingual group demonstrated a clear tendency of superior performance over the monolingual group, multivariate analyses of variance

taking the subscales into account did not statistically confirm group differences for phonological awareness in the middle or at the end of second grade. Hypothesis 1a was therefore rejected, since it could not be substantiated that the bilingual children generally performed significantly better on tasks measuring phonological awareness.

Bialystok (2002) surmised that the advantages of bilingual children in metalinguistic and phonological awareness tasks are influenced by the age of the children (likely confounded with duration of literacy instruction), the nature of the task, and the language pairs in the bilingual mix. The fact that a significant group difference was found for a scale only at the end of second grade fits well into Bialystok's summary of the phonological awareness research in which she noted a pattern of bilingual advantage in kindergarten that disappears in first grade and reappears somewhere in second grade. Essentially, she supposed that the individual differences in phonological awareness engendered by bilingual abilities are eradicated by the educational experience of formal reading instruction. This is a plausible explanation for the lack of statistical significance in this population of bilingual children with almost one and a half to two years of formal reading instruction.

Based on a series of studies that found bilingual children to have deficiencies in their L2 verbal abilities (e.g., Bialystok et al., 2005; Chiappe et al., 2002), the second sub-hypothesis regarding mean differences predicted that children in the bilingual group would perform more poorly on measures of German vocabulary than their German monolingual peers. Multivariate analyses clearly supported this hypothesis at both points of measurement. Children in the German monolingual group demonstrated higher proficiency for identifying words from pictures and finding synonyms and antonyms. In that most of the terms tested were not typical words encountered in daily urban life (e.g., "squid"), it is not surprising that they posed a substantial challenge for minority language children who had presumably been exposed to academic German language for no more than one-and-a-half years. These deficits do not have to be permanent, however. Lesaux and Siegel (2003) showed that initial deficits in L2 learners' vocabulary skills could be overcome with specific instructional strategies aimed at learning new words in the L2. They speculated that under the right instructional conditions, early deficiencies in vocabulary would not necessarily have a negative effect on reading development.

As predicted by the final sub-component of Hypothesis 1 (H1c), no significant differences were found between monolingual and bilingual participants in decoding at any point in time. This finding is in line with many other studies across a range of language combinations (e.g., Hutchinson et al., 2003; Droop & Verhoeven, 2003; Verhoeven, 2000). An explanation for this finding could be that, if vocabulary and phonological awareness are the two essential predictors of word decoding, bilingual children compensate for their weak vocabulary abilities

with their relatively strong phonological awareness abilities, thus resulting in average decoding skills. It is also probable that the relatively transparent German orthography facilitates the acquisition of decoding skills even when academic German vocabulary skills are still developing.

Although there was no formal hypothesis predicting lower reading comprehension among the Turkish-German bilingual children, it should be noted that it was unusual to find no differences in decoding or reading comprehension performance between the bilingual and monolingual readers. For example, Bos and colleagues (2003) reported substantially weaker reading comprehension skills among minority language fourth graders in Germany. Indeed, many researchers have suggested that children who have home languages other than that of the majority are at greater risk of poor literacy development when compared to their monolingual peers (e.g., Durgunoğlu, 1998; Verhoeven and Aarts, 1998). The explanation for similarities in reading performance in the current bilingual and monolingual samples is unclear, but it is probable that the participants' early stage of reading and the overall low SES of the entire sample may be factors of importance for this atypical finding.

5.3. Differential predictors of reading skills

The second hypothesis aimed to disentangle the processes involved in predicting reading and to determine the extent to which those processes differ for bilingual and monolingual beginning readers. In essence, no theories or research findings provided reason to believe that the core processes responsible for reading comprehension would differ between groups (see Fitzgerald, 1995). In other words, the literature review found no theories of L2 reading that suggested the existence of different or unique factors in the reading process compared to the L1 reading process. Therefore, sub-hypotheses 2a and 2b predicted that the same core factors would be related to word reading and reading comprehension in both groups.

The proposed predictors correlated similarly with *word decoding* performance in each group (H2a). Multiple regressions also showed similar patterns for both the monolingual and the bilingual groups. For both groups, phonological awareness was the only significant predictor of word decoding in the second and third grades. Although mid-second grade predictors did not explain a significant amount of variance in decoding abilities for the German group one year later (in mid-third grade), the end of second grade predictors explained around 15% of the variance on the decoding measure administered in mid-third grade. It is not clear why it was not possible to predict decoding for the monolingual group over a one-year period. The pattern was slightly different for the bilingual group, although it was the late second grade predictors that were not able to significantly predict decoding in third grade. Most importantly, however, for Hypothesis 2a was the simple fact that each variable made a similar contribution to the prediction of

decoding in each group. This could be interpreted as evidence that the singular predictors played similar roles for both bilingual and monolingual participants. Had a predictor been irrelevant for one group and significantly more important for the other group, Hypothesis H2a would have been rejected. This finding was directly in line with that of Chiappe and colleagues (2002) who also found that the same core components were related to word level reading for L2 and L1 readers. The fact that vocabulary skills were unrelated to word decoding among the early readers in both groups also reflected the findings of Geva and colleagues (2000) who found phonological awareness to be the only significant predictor of word decoding among a mixed group of L2 learners in Canada.

It is important to note that the amount of explained variance in word decoding was very low for both groups (between 4% and 27%). There are apparently essential factors involved in word decoding that are not accounted for by the proposed models. It is readily conceivable that, in the realm of unexplored variables, different factors would be of importance for bilingual children compared to monolingual children in predicting decoding performance. Those factors might include visual processing speed, letter recognition, or self-concept.

The *reading comprehension* skills of Turkish bilingual and German monolingual children also seemed to rely on the essentially same variables for prediction. Bivariate correlations and multiple regression analyses were used to test Hypothesis H2b, that the same core variables would predict reading comprehension for each group. The three sets of regression analyses revealed general support for the hypothesis. There were two exceptions to that finding in the analyses examining late second grade predictors for mid-third grade reading comprehension, however. First, listening comprehension played a moderate role in predicting third grade reading comprehension for the bilingual children, but no role in reading comprehension for the monolingual children. However, this difference could not be confirmed with a significant interaction effect. Second, late second grade phonological awareness appeared to be a significant predictor of mid-third grade reading comprehension for the monolingual group only. The substantial difference between the two groups with regard to the influence of late second grade phonological awareness in predicting third grade reading was statistically confirmed with a significant interaction effect. Therefore, Hypothesis 2b was supported only through second grade, in that the core predictors of reading comprehension remained the same for both groups through the end of second grade. These data indicate that, after that point, phonological awareness is no longer a meaningful predictor of reading comprehension for L2 students, but that it remains a strong predictor of reading comprehension for monolingual students. Since the core predictors of third grade reading comprehension differ for the two groups, H2b was supported only in part.

Although the same core variables were expected to predict reading comprehension in both groups, differences in the strengths of those relationships were also expected. The same analyses were used to test the final two sub-hypotheses of Hypothesis 2 and investigate the extent to which phonological awareness and vocabulary contribute to reading comprehension. Hypothesis 2c anticipated that phonological awareness would play a stronger role in predicting reading skills among monolingual children, while Hypothesis 2d predicted a stronger influence of vocabulary skills for the bilingual children on their reading comprehension abilities. Significant differences in the correlations between reading and phonological awareness were found between the two groups as well as a significant interaction effect in the regression analyses. As described above, it appeared that phonological awareness measured in mid-second grade was a good predictor of reading comprehension for both the monolingual and the bilingual students. Phonological awareness measured at the end of second grade, however, lost that predictive power for the bilingual group. Hypothesis 2c was therefore supported, but only for the differential predictive power of phonological awareness measured later in second grade. Considering the performance of the bilingual group on the phonological awareness scales approached a significant advantage over the monolingual group, it could be surmised that their metalinguistic analysis has reached capacity in the extent to which it is useful in comprehending texts. For reading comprehension, other factors become potentially more important than the perception and ability to manipulate language sounds. Therefore, the decrease in predictive power over time seems plausible and fits with the findings of several other authors (e.g., Bruck & Genesee, 1995; Campbell & Sais, 1995; Yelland et al., 1993).

The hypothesis that vocabulary skills would play a stronger role in the reading comprehension abilities of the bilingual children was not supported (H2d). It seems likely that reading comprehension at a third grade level is not yet contingent on advanced vocabulary skills. The bulk of studies that found clear and strong influences of vocabulary on reading comprehension among L2 readers were typically conducted in the fourth grade or later (e.g., Droop & Verhoeven, 2003). As Cummins and Swain (1986) suggested, it is likely that vocabulary skills will increase in importance as decoding becomes more automatic and greater levels of inference are required in advanced texts. Although the monolingual children demonstrated much stronger vocabulary skills than the bilingual children, it is probable that the level of text comprehension expected in third grade does not involve much of that more sophisticated vocabulary. In fact, Hutchinson and colleagues (2003) found increasing associations between expressive vocabulary and reading comprehension for L2 readers between the second and fourth grades. Only further longitudinal data collection with the current sample could determine if the

hypothesized differential relationship between reading comprehension and L2 vocabulary skills would emerge in later primary school when more sophisticated texts could be utilized.

In sum, the results provided some evidence for differential prediction patterns among the bilingual and monolingual children. As expected, the predictors of decoding were largely similar for both groups, and for the most part this held true for reading comprehension as well. Phonological awareness measured in the middle of second grade served as one of the most important predictors of reading comprehension for both groups. However, when measured at the end of second grade, its contribution to reading comprehension disappeared for the bilingual group. Contrary to Hypothesis 2d, this lost predictive power was not compensated for by an increasingly important role for vocabulary skills. One possible explanation could be that the reading tasks required of third graders are not yet particularly demanding with regard to vocabulary skills and that vocabulary will become increasingly important in the later primary grade levels.

5.4. Patterns of development

Unique to this study is the collection of both dual language verbal measures and written measures over several waves of data collection over several years. This allowed for a comprehensive analysis of the component processes related to growth in reading skills from the first to the third grade. Hypothesis 3 predicted that, although growth patterns would look similar for the bilingual and monolingual readers, different components of reading would influence gains in reading comprehension for the two groups. Five sub-hypotheses regarding the growth patterns and contributing components were tested to investigate potential differences between the Turkish bilingual and the German monolingual groups.

As a preliminary exploration, a set of correlation analyses explored the associations between gain scores in each longitudinally investigated scale. None of the gain scores for any of the measures correlated significantly with the gain scores from other measures. It is possible that each construct develops at a unique pace, fully unrelated to the development of the others. The lack of gain score correlations could be an indication that there is no underlying construct for growth for either group. This unexpected finding requires further examination in future research.

Since no theoretical or empirical literature indicated otherwise, it was expected that Turkish-German bilingual children and German monolingual children would demonstrate no significant differences in their growth rates on the core components of reading (phonological awareness, expressive vocabulary abilities, word decoding). This sub-hypothesis (Hypothesis 3a) was supported by a series of mixed design repeated measure analyses of covariance. Although when controlling for cognitive abilities and gender, the significant effect for time in the

phonological awareness analyses disappeared, but both vocabulary and decoding still showed significant improvement over time. In all three skill areas, the growth patterns were almost identical in the two groups; no interaction effects were found for group and time. Although similar growth rates were expected, it was surprising to see that the two groups' performance was almost identical over the four measurement points assessing decoding skills.

The similar growth rates found here are directly in line with Hutchinson and colleagues (2003) findings from a sample of mixed second-language learners in England. Hutchinson et al., however, found much larger developmental gaps between L1 and L2 readers on measures of vocabulary. It could be taken as a positive sign that formal education seemed to aid the bilingual student's vocabulary development at a similar rate to that of their monolingual classmates. It is discouraging, however, that although all other reading-related skills (phonological awareness and decoding) developed almost identically in both groups, two years of German-only instruction did not even out the bilingual group's vocabulary deficit.

As stated above, no theoretical or empirical evidence led to an expectation that L1 and L2 readers would show differential patterns in the factors responsible for their growth in word decoding abilities. After controlling for decoding at previous measurement points, stepwise regressions showed that neither vocabulary nor phonological awareness were significant predictors of growth in decoding performance. Hypothesis 3b was supported in that there were no differences between bilingual and monolingual children in the extent to which phonological awareness or vocabulary skills contributed to word decoding growth. Considering that their performance at each of the four measurement times was very similar, as were the decoding regression equations for both second and third grade, this was not surprising. Nonetheless, the fact that no significant predictors of decoding growth were found makes the evidence supporting Hypothesis 3b somewhat unsatisfactory.

No research could be located indicating that young L2 readers develop their reading comprehension skills at a different rate than their monolingual peers. Therefore, Hypothesis 3c predicted that both groups would show similar growth in reading comprehension abilities from the second to the third grade. A repeated-measures ANCOVA found no interactions between group and time, thus supporting the hypothesis.

In a similar investigation with slightly older children (third to fourth graders), Droop and Verhoeven (2003) found differential growth rates among older primary school children which was manifested in increased differences in reading performance between L1 and L2 Dutch school children over time. The authors suggested that the widening gap between the L1 and L2 readers could be explained by the children's progressive confrontation with texts containing increasingly complex and abstract language. Since the reading comprehension measures in the present

investigation contained relatively concrete language (since abstract language is not yet expected of second grade readers), it is plausible that gaps in reading comprehension development could still emerge in later primary school as reading materials become more complex. On the other hand, Droop and Verhoeven (2003), found the emerging gaps to be most pertinent when comparing L2 readers with high SES L1 readers. It is possible that within lower SES inner-city populations such as that used in the current sample, diverging patterns between L1 and L2 readers would not surface.

Similar to the hypothesis on differential predictors for reading comprehension mean *performance*, it was also expected that phonological awareness would play a weaker role in predicting *growth* in reading comprehension among the bilingual children than it would among the monolingual children. Conversely, vocabulary skills were anticipated to contribute more to growth in reading comprehension for the bilingual group than for the monolingual group. These two sub-hypotheses (Hypotheses 3d and 3e, respectively) were tested with a series of hierarchical multiple linear regressions. In effect, the analyses provided strong support for Hypothesis 3d, showing a significant interaction effect for phonological awareness and group membership in the growth equation for reading comprehension between the second and third grades. When conducting separate regressions for each group, vocabulary emerged as a significant predictor of growth in the bilingual equation only. However, no corresponding interaction effect indicating significant differences in the role of vocabulary skills was found between the two groups. Hypothesis 3e, therefore, could not be supported with tests of significance, although a trend indicating stronger predictive power for vocabulary skills was indeed detected in the bilingual group.

In sum, the analyses lead to several preliminary conclusions regarding the patterns of reading development among Turkish-German bilingual and German monolingual early readers. First and foremost, reading and base reading skills in both groups seemed to develop with striking similarity. Although in some skill areas, the time interval between measurement points was not long, even word decoding abilities measured in four measurement points over two years showed essentially identical rates of development in the two groups. It is encouraging to note that children in the bilingual group develop their German vocabulary skills at the same rate as their monolingual peers. One might expect, however, that two years of formal instruction would lead to more substantial gains in alleviating the vocabulary deficit apparent in the bilingual group (see Section 5.2. for a discussion on the groups' mean differences in German vocabulary skills). Compared to the two-year developmental lag in expressive vocabulary found by Hutchinson and colleagues (2003) among mixed L2 learners in England, the gap here is less dramatic, but the fact that the Turkish-German bilingual children in this sample remain approximately a half school

year behind their monolingual peers is still substantial. The second key conclusion obtained from these findings is that, while phonological awareness is important for developing reading comprehension from second to third grade for all children in the sample, it is less important for bilingual children. It is noteworthy that several separate multiple regression analyses in both this study and that conducted by Hutchinson and colleagues (2003) found vocabulary to be a significant predictor of reading comprehension development for L2 readers. Educators may be advised to focus particular attention on the expansion of their bilingual student's vocabularies as a method of improving the development of reading comprehension skills.

5.5. Model Fit

The final aspect under investigation represents the underlying question of each of the hypotheses: To what extent do current reading theories and, more specifically, does a model of German reading developed for early primary school children fit a sample of bilingual children? In light of the numerous international investigations pointing to large gaps in reading performance between minority and majority language children in Germany, it was important to examine how well our current understanding of reading accounts for the unique characteristics of children growing up with two languages. Hypothesis 4 aimed at investigating the fit of the Näslund and Schneider (1991) model of reading for Turkish-German bilingual children in comparison to the model fit with data from monolingual German children based on a series of structural equation models. The use of structural equation modeling allowed for the exploration of the interactions between the core variables in the proposed model on a latent basis while taking measurement error into account.

Hypotheses 4a predicted that within the proposed model, the latent construct of phonological awareness would show stronger influence on reading comprehension for the monolingual group and that German vocabulary would show stronger influence on reading comprehension for the bilingual group. Because paths were often significant for one group and non-significant for the other, it was necessary to estimate separate models for the bilingual and monolingual groups. Throughout the three sets of longitudinal analyses, the latent factor for phonological awareness consistently demonstrated stronger influence on reading comprehension in the monolingual group, while the latent factor for vocabulary skills demonstrated significant influence on reading comprehension for the bilingual group only. This clear trend provided strong support for Hypothesis 4a.

Hypothesis 4b posited poorer model fit for the bilingual group in the structural equation model for mean reading comprehension performance. Since the Näslund and Schneider (1991) model of reading was developed with monolingual children, and because bilingual children were

expected to perform differently than monolingual children on measures of the base components of literacy, it was predicted that the model would not fit a Turkish-German bilingual sample as well as a German monolingual sample. As recommended by Wicherts and Dolan (2004), the Akaike Information Criterion was used to compare model fit between the two groups. The AIC indicated better model fit for the monolingual group in both analyses in which a comparison was possible (T1 predictors of T2 reading and T1 predictors of T3 reading comprehension). In the longitudinal analysis in which a single model was found for both groups (T2 predictors of T3 reading), the explained variance for the monolingual group was almost double that of the bilingual group. This indicated that the model supplied substantially more pertinent information for predicting reading comprehension abilities for the German monolingual children than it did for the Turkish-German bilingual children. In general, although several of the variables posited by the Näslund and Schneider model proved to be insignificant predictors of reading in this sample, the key variables in the longitudinal models were able to account for around 50% of the variance in reading comprehension abilities for both groups. In one analysis, however, a great deal more variance was explained: The model using late second grade predictors to explain reading comprehension performance in mid-third grade accounted for 81% of the variance in reading performance for the monolingual group.

Finally, Hypotheses 4c and 4d provided corresponding predictions for *growth* in reading comprehension performance from second to third grade. Hypothesis 4c suggested that phonological awareness would play a stronger role for reading growth in the latent model for the German monolingual sample whereas vocabulary skills would prove to be a stronger predictor of reading growth for the bilingual sample. Hypothesis 4d predicted that the growth model would demonstrate a better fit for the German monolingual data than for the Turkish-German bilingual data. A complete model including the latent variables and manifest variables at all points in time tested these hypotheses¹.

Again, no singular model fit both the monolingual and bilingual groups due to two major discrepancies. First, the significant influence of vocabulary skills in mid-second grade on reading comprehension growth was found only in the bilingual group. Second, the direct influence of phonological awareness in late second grade on reading growth was found in the monolingual group only. These two group deviations provide strong support for Hypothesis 4c.

The large number of parameters in the growth models resulted in a substantially poorer model fit for both groups compared to the more parsimonious mean performance models

¹ It is important to reiterate here these analyses are exploratory. Accurate interpretation of the models is very difficult due to the small sample size and large number of variables in the growth models in particular.

discussed above. Neither model achieved a non-significant chi-square value and the remaining fit indices indicated a mediocre fit at best. Still, the AIC value was clearly better for the monolingual group, specifying a relatively better model fit for the monolingual group, and thus providing preliminary support for Hypothesis 4d.

The structural equation models not only reinforced the findings from the multiple regression analyses with several advantages, but they also provided some additional indications, not available from the regression analyses alone. Advantages of this approach include the ability to detect direct and indirect relationships among the variables and the additional confidence in relationships gained by accounting for measurement error. The discrepancies that did emerge between the multiple regression and the SEM analyses are likely a result of both the SEM's capacity to account for measurement error with the latent variables and ability to allow for mediation effects. The mediation effects likely shifted the regression weights and provided a more complete picture of the predictive relationships between variables.

In contrast to the regression analyses, the structural equation models more clearly revealed the unique importance of vocabulary skills for the bilingual group. This effect was no longer apparent when using vocabulary measures from late in the second grade, but for early second grade it can be concluded that vocabulary had more influence on reading comprehension for the bilingual group. It is not clear why the later measures of vocabulary are less predictive of reading comprehension. Further measurement of vocabulary would be necessary to see if that trend continues into later primary school. The literature would suggest otherwise (see Gough & Tunmer, 1986 and Proctor et al., 2005).

The inability to identify a single fitting model for both groups echoed results by Verhoeven (2000) who found that creating separate models for minority language and Dutch children produced a much better goodness-of-fit measure. Also similar to the present study, Verhoeven's LISREL analyses also found a somewhat stronger influence of vocabulary on reading comprehension for the L2 second grade readers in his sample.

It is important to reiterate that not all proposed variables influenced reading comprehension in either of the samples at hand. From the Näslund and Schneider (1991) model, one core variable, verbal memory, was left out of the structural equation models due to its negligible predictive power shown in the multiple regression analyses. Although verbal memory demonstrated substantial importance for reading in a number of studies (e.g., Brady & Shankweiler, 1991), similar to the present study, Näslund and Schneider (1991) did not find verbal memory measured in second grade to be a direct predictor of reading comprehension². It

² Näslund and Schneider (1991) found verbal memory measured in kindergarten to have a direct effect on Grade 1 reading, but not thereafter.

is therefore not surprising that no direct path from verbal memory to reading comprehension was discovered in the present sample.

In contrast to Näslund and Schneider's models, moreover, no meaningful relationships were found between vocabulary skills and verbal memory in the current data set; verbal memory only served as a predictor of phonological awareness. This is most likely explained by the fact that verbal memory in this investigation was measured on a purely pseudoword level, which was not the case in the Näslund and Schneider study³. As implied by the multiple regression analyses, integrating verbal short-term memory into the structural equations models did not change any relationships among the other core variables, nor did it increase the amount of variance accounted for in reading comprehension. Thus, for the sake of parsimony, it was omitted from the analyses⁴.

Another skill area that was measured but omitted from the final structural equation models was listening comprehension. No formal hypotheses were posited with regard to listening comprehension, but this study attempted to explore it for two reasons. First, several studies have included listening comprehension in investigating reading skills among second-language learners and found it to be relevant (e.g., Royer & Carlo, 1991). The other reason for assessing listening comprehension in this study was to diversify the measure of verbal abilities to include both expressive vocabulary measures and more context-rich listening tasks. Regrettably, the selected scale proved to have weak factor analytic structures and was only minimally associated with vocabulary and reading measures. Although the lack of predictive power of the listening comprehension measure may in part be due to its poor factor analytic structure, it may also be related to the age of the participants in the sample. As noted by Proctor et al. (2005), listening comprehension first gains importance at the upper primary school level when decoding becomes automatic and more inferential thinking is required to comprehend higher-level texts. It is likely that if the present sample were tested again in fourth or fifth grade with a well-designed listening comprehension measure, stronger predictive relationships would be found between listening comprehension and reading comprehension.

Another variable taken into account in the regression analyses but disregarded for the structure equation models due to weak associations with reading comprehension skills, was non-verbal cognitive abilities. In line with findings of many other researchers, non-verbal cognitive abilities were found to be an insignificant predictor of reading when taking skills such as

³ Although some researchers argue that any kind of verbal memory capacity is related to learning new words (e.g., Bowey, 2001; Metsala, 1999), it is ultimately not surprising that memory for pseudowords measured in this investigation was not related to vocabulary skills.

⁴ A further justification for the omission of verbal memory from the structural models was that many similar studies chose not to include it as a predictor variable at all (e.g., Proctor et al., 2005).

phonological awareness and vocabulary into account (Geva et al., 2000). However, compared to the Näslund and Schneider German reading research (1991, 1996; Schneider & Näslund, 1999a, 1999b, 1993) in which similar analyses were conducted with a comparable set of data, non-verbal cognitive abilities was a substantially weaker predictor in this study. Näslund and Schneider found IQ measured in preschool to explain around 17% of the variance in reading comprehension in second grade, whereas in this study, cognitive abilities measured in mid-first grade accounted for no more than 2% or 9% of the total variance in reading comprehension in second or third grade for the bilingual and monolingual children respectively⁵. Additional research with similar samples is badly needed in order to better understand the role of non-verbal IQ in the reading development of L2 readers.

Overall, general support was found for many of the proposed hypotheses. Although significant differences were found between the two groups on measures of vocabulary and, as hypothesized, phonological awareness was a more important factor in early reading for the monolingual group, it can essentially be concluded that there are more similarities than differences between the German monolingual children and the Turkish bilingual children in their base reading skills and literacy development. The high level of similarity between the two groups on the German reading measures was particularly surprising. The most plausible explanations for that seems to be the early stage of reading during which the sample was investigated (in which demands on advanced German vocabulary knowledge were still minimal) and the similarly low socio-economic status of the children's families in both groups. Nonetheless, as expected, the theoretical model explained less variance for the bilingual group than for the monolingual group; although neither group fit the proposed theoretical model in its original form. In sum, the findings discussed in this section provided multiple noteworthy insights into the development of reading skills among Turkish-German bilingual children compared to their monolingual peers. The results of this investigation, however, must be taken as preliminary due to several restrictions of the study design and sample. A discussion of those limitations and applications for the findings follows below.

⁵ Although the source of this minor discrepancy between the present findings and the Näslund and Schneider (1991) is unclear, there are several noteworthy differences between the two studies that may have contributed to it. First, the Näslund and Schneider study was conducted with volunteer children from Munich who were primarily from households of average SES. In contrast, the current sample was drawn from urban Berlin districts with very low SES. Furthermore, the children in the Munich sample were volunteered by their parents, whereas the Berlin sample here was selected as a sub-sample from a larger school system investigation in which parent data or permission was not collected. Second, the utilized measures of cognitive abilities differed between the two studies. Third, the Munich study was conducted from 1984 to 1993. Since the collection of data for the first and second graders in this study took place two decades after that of the Munich study, cohort differences could be a further viable explanation for the different findings in the two studies. All three of these differences between the two studies could affect the role of non-verbal cognitive abilities in the process of acquiring reading skills.

5.6. Limitations

As a first investigation of early reading processes among bilingual Turkish-German children, this study includes a broad set of robust reading measures over a period of two and a half years. Nonetheless, due to a lack of resources and the exploratory nature of the study, there are several shortcomings that require mention. The aim of this section is to help future research or replications of this work avoid several of the obstacles met in this investigation.

The most regrettable limitation of this investigation was the inability to gauge the socioeconomic or educational status of the participants' parents. There is some research indicating that SES has little influence on narrow components of literacy (such as phonological skills; Samuelsson & Lundberg, 2003). Nonetheless, SES is an important control variable and would have been useful for verifying the comparability of the Turkish-German bilingual and German monolingual groups. Measures of SES are key components of many well-validated models across a wide range of academic competencies for broad populations internationally (see Watermann & Baumert, 2006). According to a study by Ransdell and Wengelin (2003), SES was as important, if not more, than bilingualism in determining the linguistic skills related to literacy in bilingual children. They found the occupational status of bilingual children's mothers to be one of strongest predictors of the children's writing quality.

Due to a lack of personnel and financial resources, it was not possible to directly ask the participants' parents about their educational levels or occupations. Both in the overarching BeLesen study and in this embedded investigation, the attempt was made to overcome the lack of parental questionnaires. As part of the BeLesen study, participants' teachers were asked to estimate the income levels of the participants' parents. Many teachers left this item unanswered, presumably because they did not know or feel comfortable providing such information. The current sub-investigation also attempted to find indicators for home SES. In mid-second grade and late second grade, the children in this sub-sample were asked to indicate how many books they have in their homes, a common and highly predictive measure of parents' educational backgrounds (e.g., Lehmann & Nikolova, 2005). Although self-report instruments of this nature have demonstrated reliability among high school students and primary school children as young as fourth grade, the current data indicated that second grade children are not capable of reliably assessing the books in their homes (correlations between T1 and T2 did not exceed .50 for either group). Despite best efforts to find indicators of SES without access to parent reports, neither of these attempts was particularly satisfactory. The only remaining indicator of SES available was the SES categorization of the zones in which the participating schools were located. The school district is a good estimate of the type of neighborhood in which the families lived (all school districts in the study were ranked as socio-economically disadvantageous or highly

disadvantageous). Since the groups were sampled equally from each school district, they did not differ with regard to the SES ranking of their school districts, thus allowing for some confidence in the socio-economic comparability of the two groups. Additional parent information would have been valuable in confirming this assumption.

Because the children in this sample were investigated only during the earliest stages of literacy, during which reading requires relatively simple vocabulary and contexts, the available data cannot be used to examine the consequences of lower L2 vocabulary skills for more demanding reading materials. Theorists and researchers suggest that L2 verbal and vocabulary abilities become increasingly important, as decoding is mastered and reading processes shift toward requiring greater levels of inference in context-reduced texts (e.g., Cummins & Swain, 1986, Proctor et al., 2005; Schneider, 2004). It therefore seems that the expected higher levels of predictive power for German vocabulary skills among bilingual children would more likely be found in later primary school. It is unfortunate that this study could not extend the longitudinal design to include data collection through the fifth or sixth grade. It would be important to observe possible changes in the gap in vocabulary skills between the two groups and the later effects of vocabulary skills on more advanced reading comprehension abilities.

The lack of a further language comparison group makes it impossible to disentangle the effects of Turkish-German bilingualism and bilingualism itself. The phonological structure of the Turkish language with its simple syllable structure, vowel harmony, and post-inflection system led to the hypothesis that phonological awareness would be more advanced in a population of Turkish-speakers than in a population of German-speakers whose language does not require the same levels of phonological perception. The theories regarding increased phonological awareness among bilinguals strengthened the hypothesis that Turkish-German bilingual children would have heightened levels of phonological awareness. However, in order to demonstrate that the phonological abilities of the bilingual group were indeed shaped by their experiences with speaking Turkish at a young age, the inclusion of further bilingual groups would have been necessary. It was therefore not possible to separate the potentially differential effects of bilingualism from the influence of Turkish abilities on phonological awareness. The necessity of understanding the effects of language specificity and bilingualism on phonological awareness is diminished, however, by the fact that no significant differences between German monolinguals and Turkish-German bilinguals were found at all. Since only a tendency was found within this investigation, it is not essential to explain the root of a bilingual advantage for phonological awareness.

Although it could be argued that this study lacks generalizability, it should again be noted that this specific population was intentionally selected for several reasons. First, a great deal of

research has demonstrated the impact of specific linguistic backgrounds on phonological awareness (e.g., Bialystok et al., 2005), thus underscoring the importance of examining each linguistic group individually, not over a broad range of language combinations. The use of a mixed sample of L2 learners might be more easily generalizable for a broad range of L2 learners, but would have led to minuscule cell sizes for each linguistic combination and an overall loss of statistical power and precision. Second, the Commission on Behavioral and Social Sciences and Education National Research Council specifically recommended generating research on additional language groups in their 1997 research agenda for improving schooling for language-minority students (August & Hakuta, 1997). As the first investigation on the reading processes of bilingual Turkish-German children, this study fills that recommendation. Third, because populations of Turkish children make up the largest proportion of minority students in Germany, this group was deemed most important for research in the area of L2 learners and literacy in Germany. Therefore, this sample was not collected with the aim of being representative for all German students or all minority language students in Germany; on the contrary, each linguistic group likely has specific unique characteristics and needs.

The final notable limitation in this investigation is the common predicament of sample size. Although a power analysis (described in Section 3.2.1.) indicated that the group sizes were large enough for the fundamental analyses planned for the investigation, somewhat larger samples are required for structural equation modeling. Taken together, the entire sample of 169 could have been adequate for interpreting the results as solid findings. But since the bilingual and monolingual groups could not be analyzed together with a single model, separate AMOS analyses had to be run for each group with quite small sample sizes. The structural equation models presented in this paper can therefore only be considered exploratory. Nonetheless, the SEM results are good reflections and extensions of the findings from the multiple regression analyses, thus lending them substantial credibility. It is not uncommon for research of this nature, in which in-depth measures must be collected individually with each participant and over a several points of measurement, to include exploratory SEM analyses with smaller sample sizes (e.g., Verhoeven, 2000). Ultimately, limited resources did not allow for the collection of more data in individual testing sessions as was necessary for the verbal measures in this study.

Overall, this investigation could have been strengthened had more resources been available for several extensions including parent questionnaires for collecting background SES information, a longer data collection period extending through the later primary grades, an additional bilingual language group, and an overall larger sample size. Any replications of this research would be well served to address these empirical deficits. Additional recommendations for future research are discussed in the following section.

In spite of the investigation's limitations, the findings provide several important implications and a strong impetus for further research. The results of this investigation showed that, with the exception of German vocabulary knowledge, a sample of Turkish bilingual and German monolingual inner-city children perform similarly on all measures of reading and base reading skills through the third grade. Not only does this study demonstrate the equal potential of Turkish-German and German children at the onset of primary school, but it also highlights the need for more work on developing reading models that better explain the essential factors involved in reading for bilingual children. This investigation provides the first longitudinal examination of how several aspects of literacy develop among a group of Turkish-German bilingual children in the German school system.

5.7. Recommendations and implications

From both the findings and limitations of this investigation, a number of recommendations for practice and research can be drawn. The deficient state of research on literacy acquisition among minority language children in Germany leaves a great deal of room for new studies, for which this report provides a solid starting point. Furthermore, the current movement in Germany toward initiating educational programming aimed at boosting language and literacy skills among minority language children provides fertile ground for utilizing findings from the current study. Suggestions for research and practical extensions of select findings from this investigation are discussed in detail in this section. Both empirical and intervention studies are clearly needed to address the many questions left unanswered and the substantial difficulties faced by young Turkish-German schoolchildren.

5.7.1. Recommendations for future research

The first recommendations stem from aspects of the current research project that could not be addressed for this study, but should be taken into account in future studies. For example, as discussed above, since phonological awareness has demonstrated language-specific qualities (e.g., Caravolas & Bruck, 1993), it would be useful to examine the phonological capabilities of children with varied bilingual language combinations. In this case, it could help partial out the potential effects of bilingualism itself and the potential effects of Turkish abilities on phonological awareness if future research were to include one or two further comparison groups from languages with differing phonological structures (i.e., Russian/German or Vietnamese/German).

A second recommendation for future research involves the socio-economic status of the participants. Although the lower SES background of the bilingual children in this sample is typical of minority language children in Germany, it is plausible that the current results are only

applicable for inner-city Turkish-German children. It would be useful to examine groups of bilingual and monolingual children in higher SES contexts. In order to ensure that the current findings are indeed related to Turkish-German bilingualism, and not restricted to Turkish-German children in inner-city populations, similar research should be conducted with Turkish-German bilingual children in other socio-economic contexts. Augmenting samples with corresponding bilingual children in the counterpart country could be another way for future researchers to add depth to their findings on the effects of bilingualism on reading. In this case, the replication of these results with a sample of Turkish-German bilingual children in Turkey could demonstrate the generalizability of these results with regard to the effects of Turkish-German bilingualism, regardless of cultural or socio-economic context.

The extensive and diverse models of reading provide multiple aspects of reading and base reading skills that could or should be included in a well-designed study of reading comprehension. In light of the substantially lower amount of variance explained by the simple theoretical model for the bilingual group, it is recommended that in future research on models of bilingual reading development, a broader range of instruments is used over a longer period of time. Ultimately, the factors thought to underlie literacy development in beginning readers did not account for the same amount of reading performance in the bilingual sample as they did in a monolingual sample. Studies similar to the nine-year longitudinal study in Munich (Schneider & Näsland, 1993) would be invaluable for monitoring the academic and linguistic development of second-language learners with a wide range of instruments, from the first institutional exposure to the second language in a preschool or kindergarten to a level of advanced reading comprehension in late primary school. Measures of grammar, pre-literacy skills such as alphabet knowledge, family reading practices, and non-word decoding should be included whenever possible. It is important that the additional variables contributing to reading abilities are discovered and that appropriate models of reading are established for multilingual populations. Only when the base components at work are understood can the optimal pedagogical approaches to literacy instruction in multilingual classrooms be developed.

5.7.2. Implications for educational practices and policy

The present findings regarding phonological awareness provide a basis for considering the potential benefits of phonological awareness training for Turkish-German children in early primary school. In that the bilingual and monolingual groups performed and developed similarly on measures of phonological awareness, it is reasonable to predict similar enrichment potential from intervention programs aimed at phonological awareness skills. The bulk of intervention studies using phonological awareness training programs have produced convincing evidence

regarding their effectiveness in boosting early reading skills for both L1 readers (e.g., Schneider et al., 1997) and L2 readers (e.g., Lesaux & Siegel, 2003; Stuart, 1999).

Among monolingual children, phonological awareness training (usually accompanied by teaching the representation of speech sounds with letters) has demonstrated relatively consistent beneficial effects on reading development. A meta-analysis of phonological awareness training studies found moderate to large effect sizes for reading skills in both randomized and matched designs ($d \geq .70$; Bus & IJzendoorn, 1999). Although the long-term effect sizes were smaller, the implications for opportunities to help children improve reading skills through phonological training are important. It is possible that introducing reading through a program grounded in phonological awareness skills and letter-sound correspondences could capitalize on the metalinguistic strengths of second-language learners, whereas the more holistic approaches emphasizing word meaning are based on second-language learners' potentially weakest area (Stuart, 1999).

An exemplary training study by Stuart (1999) involving 112 five-year-olds from inner-city schools, 86% of whom were L2 learners, suggested that L2 readers can gain an advantage in literacy acquisition if phonological awareness training is provided. Half of the randomly assigned participants were given a 12-week intervention of phoneme awareness and phonics training. This training, easily provided by classroom teachers, accelerated the children's acquisition of phoneme awareness and phonics knowledge, as well as their ability to apply those skills in reading and writing. Inner-city and L2 learning children who had participated in the phonetic training experimental group improved significantly on measures of reading accuracy and spelling in comparison to the control group, with effects lasting one year following the intervention. The experimental group also made substantial gains on measures of reading comprehension that were marginally significant ($p < .06$) compared to the control group. This study demonstrates that early structured teaching of phoneme segmentation accelerates the development of phonological and reading skills in five-year-olds, including children learning a second language.

A German program similar to the 1999 Stuart intervention study is currently under evaluation by researchers from the University of Würzburg with a sample including a substantial proportion of minority language children. The interim results seem to indicate that phonological awareness programs benefit L2 learners' literacy development. Still, following the intervention, the L2 learners scored more poorly on the German-language based phonological awareness and reading measures than their German monolingual peers (Weber, Marx, & Schneider, 2005).

One of the clearest findings of the current study was the pervasive delay in German vocabulary skills demonstrated by children in the Turkish bilingual group. This was an expected finding. It is commonly reported that L2 learners lag behind their L1 peers with regard to

vocabulary development in the language of instruction (e.g., Verhoeven, 2000). In line with the strong recommendations of Proctor et al. (2005) as well as Chiappe et al. (2002), the results of this investigation serve as a evident indication that there is a need for interventions aimed at L2 vocabulary development within the German school system. In his similar study with minority language children in the Netherlands, Verhoeven (2000) clearly interprets the lack of L2 vocabulary and oral skills as an indication that L2 vocabulary growth should be stimulated in preschool and kindergarten children.

A host of well-evaluated intervention programs demonstrates substantial potential for vocabulary enrichment, some of which are quite simple to implement. Several studies have shown that reading aloud and discussing of stories with pre-readers and early readers can lead to increased vocabulary knowledge. Morrow (1992) and Morrow and Smith (1990) have shown interactive story reading in small groups to be a particularly good method for teaching young L2 learners. Robbins and Ehri (1994) reported the results of an experiment in which pre-reading children were read a story twice several days apart, then given a vocabulary test with a selection of new words in the story. They found that simply reading the same challenging story twice within a week increased vocabulary of the young participants. A recent German intervention study conducted by McElvany and Artelt (2005) demonstrated that a four-month parent-child reading program significantly increased vocabulary knowledge among early primary school children. Although gains were somewhat smaller for children with poorer initial vocabulary scores, children from families with lower levels of oral communication skills showed the greatest benefits from the intervention program with regard to their use of elaborate verbal communication. Similarly, Golova et al. (1999) and High et al. (2000) conducted a randomized controlled reading intervention study with Spanish-speaking parents of toddlers. They not only found exceptionally high compliance and an overall increase in literacy practices in the home, but also significantly higher expressive vocabulary scores among toddlers ages 18-25 months, thus indicating the viability and effectiveness of such reading interventions with minority language families. Carlo et al. (2004) conducted one of the few such quasi-experimental studies with a small sample of Spanish-English primary school children throughout the United States. The authors found that 15 weeks of vocabulary training significantly improved the vocabularies and reading comprehension performance of the fifth grade bilingual students compared to both L2 and L1 control groups. The bilingual group experienced equal gains as the monolingual group as well.

Programs such as these would be easy to integrate into mainstream classroom practices, but must be evaluated with regard to their effectiveness for L2 learners in Germany. Although the poor vocabulary knowledge of the current sample did not seem to have a negative effect on students' reading skills, it is likely that, as the texts with which they are presented in later grades

become more demanding, their deficits in vocabulary knowledge will have detrimental effects on their academic progression across a range of academic subjects. German literacy intervention studies such as those described above aimed at improving vocabulary knowledge and acquisition strategies for minority language children are strongly recommended.

Based on existing intervention studies and the findings of this investigation, it is likely that that phonological awareness training would be similarly effective for both bilingual and monolingual children. On the other hand, the literature and the present findings suggest that bilingual children would particularly benefit from interventions and pedagogical programs aimed specifically at boosting L2 vocabulary skills. To better understand and service the growing population of multilingual children in Germany, a great deal more intervention research of this nature is needed across a wide range of skill areas, particularly with consideration for methods practicable in the mainstream classroom.

The final recommendation derived from the results of this study underscores the need for empirical investigations of language-based interventions through kindergarten and preschool programs geared toward second-language learners in Germany. Much of the international educational literature today places great emphasis on the necessity of beginning early with well-founded, pedagogically comprehensive language support for at-risk and minority language children (e.g., Grimley & Bennet, 2000).

Through a review of several empirically-thorough longitudinal studies of oral language and reading, Hagtvet (1993) concluded that at-risk children may develop their academic and linguistic skills more erratically than typical children. Hagtvet proposes that this inconsistent development makes at-risk children more susceptible to negative environmental influences than children with superior socio-economic or linguistic resources. However, Hagtvet also argues that that susceptibility also applies to positive influences such as interventions, provided they take place early (before grade three). She found that the integration of language and phonetically rich components into kindergarten curriculum enabled at-risk children to learn at a much higher rate than would otherwise be expected in the first grade.

The well-documented benefits of fostering early childhood linguistic and academic development has been largely disregarded in Germany thus far. According to a review conducted by Spiess and colleagues (2003), the very limited availability of early childhood programs in Germany leaves most parents only with the option of arranging private care for their young children. Furthermore, the authors point out that surprisingly little German research has examined the impact of early childhood programs on later academic performance. In their analyses of a longitudinal German data set, Spiess and her colleagues discovered that, indeed, attending early childhood programs significantly raises the probability of immigrant children

gaining entry into higher levels of the German educational system at the commencement of primary school. This was not the case for native German children. The authors draw the conclusion that the opportunity for immigrant children to acquire German language skills is the central reason that kindergarten attendance is the strongest predictor of later admission into higher levels of schooling. Finally, Spiess and colleagues (2003) call for more political attention to the need for early institutional pedagogical (and thus linguistic) support for children from immigrant families. The deficits in the verbal abilities of L2 learning children found in the present investigation only underscore the validity of this suggestion. According to the BeLesen data set utilized for the present study, most inner-city Berlin children from immigrant families do attend some type of preschool, however, these programs presumably lack pedagogical content aimed at the specific needs of minority language children. Studies examining the long-term effects of general preschool attendance compared to the effects of preschool programs with systematic language support would be valuable for partialling out the effects of the two approaches to early learning.

Researchers have noted for many years that a large segment of children in Germany is impeded in its academic development by the lack of early language support through pedagogically-based and easily-accessible preschool programs. Nearly three decades ago, Rist (1978) called for early intervention research, explaining that simply living in Germany is not adequate for learning the levels of German language required for academic success.

To be in German society, but not of it, does not seem to be a sufficient basis upon which to assume the children are equipped to perform in German classes. The reason for stressing this particular point is that the preschool program could fill a vital need as a transition period for the children from the language and milieu of the home to the language and expectations of the German classroom[...] While neither preschool nor kindergarten is compulsory in Berlin, it would be of interest to speculate on the impact a program more attuned to the guest worker children might have on their later school careers. (pp. 241-242)

It is unclear why this has not been the topic of more research or action in the last 30 years. There is, however, reason to believe that recent strong recommendations for changes in Germany's early education programs from experts such as the United Nations Special Rapporteur on the right to education (see Peter, 2006) and the OECD (2006) are slowly being taken more seriously.

The current investigation points to the importance of phonological awareness and vocabulary skills in the processes of early reading development for both bilingual and monolingual students. For both skill areas, numerous intervention models exist that can be integrated into classroom instruction or utilized in preschool settings. This study reflects other research indicating that L2 vocabulary-based programs may be the most valuable interventions for bilingual children. Although the benefits of early childhood support for minority language

children are well documented in the literature, they have not yet been harnessed to their full potential in Germany. In sum, there is a great deal of information already available in the literature; the task at hand is to assess the efficacy of the available programs in German preschool and elementary classroom settings for children acquiring literacy in second languages.

5.8. Conclusion

Reading and comprehending German written language are vital skills for minority language children not only within the German academic system, but in society in general. Being a skilled reader is essential for gaining access to educational opportunities and the benefits of the larger society (see the definition of reading literacy in Baumert et al., 2003). This understanding of reading comprehension makes clear the discouraging implications of the international comparative school studies such as PISA and IGLU, in which it has been clearly shown that no other country leaves children of minority language families further behind in their reading competencies than Germany (see Schwippert et al., 2003). Interestingly, in this sample of particularly young inner-city children, no discrepancies in reading abilities were found between minority language children and first language German speakers. However, by exposing the lack of a well-fitting model of reading for second-language learners, this study suggests that the current state of our empirical and theoretical understanding of reading processes is not serving the growing proportion of minority language students in Germany as well as it should. The increasing heterogeneity of the German schools and classrooms requires educators and researchers alike to ensure that research, theories, and instructional practices are conceptualized with consideration for children of diverse linguistic backgrounds.

Although heterogeneity is increasing in the German school system, the question of integrating children of immigrant families has been discussed at the periphery for several decades. It is worthwhile to put the findings of this study into an historical perspective. In 1978, a visiting scholar of the Max Planck Institute for Human Development wrote a book on the condition of “guestworker” children in Berlin (Rist, 1978). The parallels in his findings thirty years ago to the current socio-educational landscape are striking:

... [It] does not seem justifiable that the Berlin schools can proceed with no adjustments on their part. . . the schools can hardly refuse to take [the guestworker] constituency into account. . . The essential structures of German education survived the Nazi period, survived the efforts at reform instigated by the Allies in the postwar period, survived the efforts at change from the protest movement of the 1960s, and is now not about to budge on account of foreign children (p. 237)

There is little evidence that substantial change has occurred since that the time of Rist’s analysis of the educational opportunities for minority language children in the 1970s.

It has been thirty years since the U.S. abandoned the “sink or swim” approach to education for second-language learners. Since the 1970s, educators have tested a wide range of educational approaches for meeting the needs of those minority language children. The new methods have been aimed at helping students develop proficiency in English as well as learn the knowledge, skills, and attitudes that make up the curriculum (August & Hakuta, 1997). In response to increased interest in instructional methods for L2 learners, the scientific community also began developing a strong research base on L2 learners with special regard for language and cognition. These events have resulted in a “rich portfolio of research on English language learners, ranging from basic processes to program evaluations and from program characteristics research to the collection of national statistics” (August & Hakuta, 1997, p. 2). This is not to say that U.S. researchers have perfected instructional methods for L2 learners; that is not the case (see Carlo et al., 2004; Limbird & Stanat, 2006). Still, their 30 years of research can provide educational scientists with sufficient information for initiating investigations aimed at understanding the unique needs and abilities of minority language children in Germany. It is anticipated that this study will be one of many examining the linguistic, cognitive, and academic processes involved in the education of minority language students in Germany.

This research project was conceptualized during a time in Germany when second-language learners and the educational system in which they were integrated was receiving an unprecedented amount of attention on both political and educational systems levels. This study of bilingualism and literacy acquisition attempted to contribute to building a foundation of empirical German research for understanding the linguistic and literacy development of those minority language students. Grounded in international research and theories on bilingualism in conjunction with a current German model of reading comprehension, this study made a first attempt to understand the processes involved in learning to read in German as a second language among native Turkish speakers. For the findings to be as relevant as possible for the current educational discussion of policies for integrating and best serving minority language populations, this study investigated what would be considered a high-risk population of inner-city early primary school children in schools with large percentages of non-native German-speaking students and neighborhoods with low socio-economic status. Also to enhance applicability, this study asked two essential questions, pertinent for the current political and educational discourse: 1) Do Turkish-German bilingual children have any linguistic advantages over German monolingual children that can be utilized for teaching early reading skills, and 2) Does the German reading development process differ for the Turkish-German children compared to their monolingual classmates?

With regard to the first question, contrary to predictions based on theories and findings in the bilingualism literature, the bilingual children in this study demonstrated only a tendency to out-perform their monolingual peers on measures of phonological awareness, producing few significant differences; while their performance on measures of German vocabulary knowledge was clearly deficient compared to German monolingual classmates. Still, the bilingual children demonstrated reading skills identical to their native German-speaking peers at each point of measurement. There are many possible explanations for that finding, including the presence of enhanced skills in additional areas of metalinguistic awareness or further unidentified abilities that help the Turkish-German children adapt to reading in a still developing second language. This investigation does not conclusively answer the question if Turkish-German children have phonological advantages over their monolingual peers, but it does demonstrate the ability of the young bilingual children in this sample to compensate for the lack of vocabulary knowledge in their second language. More research is needed to determine what special skills are at work in this compensation process.

It is important to recognize that in addition to the ability to compensate for a still developing vocabulary, the bilingual children possessed the ability to communicate in two languages, a definitive linguistic advantage on a larger scale. Bialystok (2002) draws a similar conclusion from her review of 25 years of bilingualism research: “The differences between monolinguals and bilinguals that occur are invariably to the benefit of the bilinguals. Knowing more has never been a disadvantage when compared to knowing less” (p. 192).

The second essential question in this study addressed the issue of developing reading skills. Findings showed that although similar base components play a role in learning to read for both bilingual and monolingual children, these components manifested themselves differently for the two groups, with phonological awareness often demonstrating a stronger effect on reading comprehension for the monolingual group, and vocabulary knowledge playing a stronger role for the bilingual group. These divergent patterns can only currently be described as tendencies, however. The greater proportion of unexplained variance and poorer fit of the models for the bilingual children clearly exposes the need for more research with larger samples and a wider variety of instruments.

It should not be forgotten that the Turkish-German children in this sample were born in Germany, possessed German citizenship, had by and large attended German preschools and daycare before coming to school, and were integrated in the German school system their entire lives. Although this may be apparent in the equal reading scores of the two groups under investigation, it is not apparent in the substantial gaps in vocabulary knowledge between the German monolingual and Turkish-German bilingual children who had experienced the same

amount of education with the same group of teachers. If German educators choose to accept the challenge of a multi-cultural society, then it is time for the scientific community to begin researching the processes at work in second language literacy acquisition, devising effective interventions and instructional strategies, and educating the educational community about their findings.