

17

PhD Revisited: Young language learners

The acquisition of English in Norwegian first-grade classrooms¹

ANNE DAHL

NTNU – Norwegian University of Science and Technology

ABSTRACT This chapter reports a PhD study (Dahl, 2014) on the effects of English target language input in Norwegian first-grade classrooms, comparing three learner groups with different volumes of input. The results highlight the role of input for acquisition, showing that with sufficient exposure, development in a foreign language can be rapid and similar to other forms of second language acquisition. Overall contributions and practical implications of the study to the field of English teaching are presented.

KEYWORDS second language acquisition | foreign language learning | early start | age of acquisition | target language input

-
1. The chapter presents the overall results of a PhD study (Dahl, 2014) from NTNU Norwegian University of Science and Technology. This is an article-based thesis with two published articles (Dahl, 2015; Dahl & Vulchanova, 2014). The PhD thesis in its entirety can be found here: <http://hdl.handle.net/11250/244422>.

INTRODUCTION

An apparent contradiction in language acquisition research is between findings indicating that an early age of acquisition (AoA) for language is normally beneficial for ultimate attainment, and those indicating that an early start with a foreign language in school does not necessarily entail better end results than a later start. The explanation for these conflicting results may be a lack of input in many such classrooms. Since the main benefit of a young starting age has been found in the ability to acquire language implicitly from the input, not from explicit instruction, the observed lack of effects of an early start in a new language at school may be a result of the learning situation being more suitable for older learners. This was the overall hypothesis of the present PhD study.

The main research questions can be summed up as follows:

To what extent and in what ways do different volumes of target language exposure lead to increased language competence in an early-start L2 classroom?

THEORY

An important theoretical question in language acquisition research is the role of age. Research findings imply that there are differences in how young children acquire language compared to older children and adults. With exposure to multiple languages from early childhood, native competence in each language is typical. On the other hand, such competence is extremely rare, if at all possible, in a language acquired after early childhood (e.g., Hyltenstam & Abrahamsson, 2003). What constitutes *early childhood* here is not clear. Original proposals of a *critical period* for language acquisition placed the age limit for native-like acquisition at puberty (cf. Lenneberg, 1967; Penfield & Roberts, 1959). However, later proposals outline a more complex picture. Changes in our language-learning capacities seem to start much earlier, and no clear cut-off point for these capacities has been found; there may in fact be multiple smaller periods in which humans are more sensitive to various aspects of language input, or the decline in language learning abilities may even be linear and start at or soon after birth (see for example Hyltenstam & Abrahamsson, 2003, for overview and discussion).

The causes of our decline in language acquisition abilities with age are also not clear. Suggestions include lateralization or loss of brain plasticity (cf. de Bot, 2006; Lenneberg, 1967; Penfield & Roberts, 1959; Pulvermuller & Schumann, 1994), or social factors such as inherently better motivation in children, or the

input being more substantial and more appropriate for children (e.g., Bialystok & Hakuta, 1999; Piske & Young-Scholten, 2009). Others argue that adults' use of cognitive problem-solving mechanisms during the learning process inhibit acquisition, or that acquisition is hindered by affective factors in older learners (Krashen, Long, & Scarcella, 1982; Rosansky, 1975). N.C. Ellis (2006) argues that factors shaped by the L1 hinder L2 acquisition. "Entrenchment" of the L1 system with use in such an account would also result in age effects.

REVIEW

Evidence for the advantage of an early AoA of a new language has been abundant in studies within the field of second language acquisition (SLA) (see for example Hyltenstam & Abrahamsson, 2003; Muñoz & Singleton, 2011). However, the advantage has been found mainly in the end result; ultimate attainment is more likely to be successful for younger learners, but younger learners are not generally faster than older learners (e.g., Krashen, Long, & Scarcella, 1979; Snow & Hoefnagel-Höhle, 1977). Still, with the knowledge that younger learners normally end up with higher levels of L2 competence than older learners, it has seemed intuitively wise to lower starting ages for foreign languages in school.

However, at the time of the PhD study revisited in this chapter, a general finding in studies of the effect of early foreign-language instruction in formal settings was that there is no benefit of an early start, at least in the long run. This was the case in early studies of early-start English in Japan (Oller & Nagato, 1974), French in Britain (Burstall, 1975), and English in Sweden (Holmstrand, 1982). A number of later studies pointed in the same direction, finding no advantage for younger learners, and in some cases even that instructed settings favored later starters (see for example contributions in García Mayo & García Lecumberri, 2003; and in Muñoz, 2006).

On the other hand, in addition to differences in ultimate attainment, younger learners had been found to acquire language differently than older learners and to depend more on substantial input in the target language (see for example Murphy, 2010, for an overview of research). Specifically, findings implied that young children learn implicitly from exposure while explicit instruction and formal learning may be more useful for older learners. Thus, the lack of advantage found for an early start with foreign languages may have been a result of a lack of target language input in such contexts. The question was thus whether foreign language learning could be said to constitute a qualitatively different process such that the advantages of a young AoA would hold for naturalistic second language acquisi-

tion only, or whether foreign language learning might in essence be seen as a type of SLA, where increased exposure to the target language could be effective.

METHODOLOGY

This PhD study investigated English acquisition in first-grade students in different learning contexts in Norwegian schools, where volume of input in English was the main difference.

PARTICIPANTS

Participants in the study were a total of 82 students in four groups:

A group of monolingual Norwegian students ($n = 29$, mean age at pre-test 6;1) in a Norwegian state school, receiving regular English instruction as mandated by the curriculum (Utdanningsdirektoratet, 2006). The students all attended the same school, but different classes, and a number of teachers were involved in English teaching throughout the school year. Since English teaching in this group mainly relied on Norwegian as the language of instruction, this group is referred to as the “L1-based group” in this chapter.

A group of monolingual Norwegian students ($n = 31$, mean age at pre-test 6;1) in a Norwegian state school, receiving normal English instruction as mandated by the curriculum, but, crucially, where teachers focused heavily on providing English input both inside and outside English class throughout the school day. These students all attended different classes in the same school, which was not the same school as that of the L1-based group. This school was selected because teachers were willing to provide more substantial input in English than what is common. Again, a number of teachers were involved in English teaching; one was a native speaker of English, but she also spoke and taught other subjects in Norwegian. L1 Norwegian was not excluded in English classes in this school, and this group is thus referred to as the “bilingually based group” in this chapter. The two state schools were situated in similar neighborhoods in terms of socio-economic factors, and had previously scored similarly on national tests of English.

A group of students ($n = 7$; mean age at pre-test 6;0) in international schools in Norway, where English is the language of instruction, starting school without competence in English. They were selected based on reports from parents and teachers that they did not know English upon starting school. One of these students was a monolingual speaker of Norwegian, five were bilingual speakers of

Norwegian and another European language, while the third was trilingual before learning English in school. This group is referred to as the “immersion group”.

A group of students ($n = 15$; mean age 6;6) in international schools in Norway who were native speakers of English and served as a control group in the study. Most of these students also had competence in one or more other languages, usually Norwegian, but their parents classified English as a native language for all of them. These students only completed the tests of English proficiency for which no norming existed; see table 17.1 below. This group is referred to as the “control group”.

RESEARCH DESIGN

The research design consisted of quantitative comparison of the three learner groups (excluding the control group) in a pre-test upon starting school and then again after the first year on a number of measures of English competence. Data collection consisted of a pre- and a post-test session with each participant, each lasting approximately one hour. A number of background and outcome measures of verbal and non-verbal ability and comprehension of English were used to measure English development from the beginning of the school year, approximately in September, to the end of the school year, approximately in May. The English native speakers in the control group were tested once, approximately halfway through the school year, and only on English measures, excluding English vocabulary, since the test used for this measure has been normed for native speakers, i.e. its scores can be systematically compared to what has been found in a large, representative group of English speakers in the USA. Table 17.1 gives an overview of the tests used in the study.

Teaching methods in the classrooms were not carefully controlled, and no lesson analysis was performed, but teachers reported on classroom practices, materials, and time spent on English throughout the year, in particular for the two state-school groups that were otherwise very similar. In the international school immersion group, information about factors such as whether students had attended English as an Additional Language (EAL) classes was collected, but teaching methods were not analyzed.

TABLE 17.1. An overview of tests used in the PhD study.

Background measures		Outcome measures	
Test data	Analysis	Data	Analysis
PPVT-IV form B	<ul style="list-style-type: none"> ▶ Growth Scale Values (GSVs) ▶ Age equivalents ▶ Within-group and between group comparisons of differences in means ▶ Multivariate analysis of co-variance (MAN-COVA) comparing groups ▶ Correlation analysis with other measures 	PPVT-IV form A PPVT-IV form B	<ul style="list-style-type: none"> ▶ Growth Scale Values (GSVs) ▶ Age equivalents ▶ Between-group comparisons of differences in means ▶ Within-group comparisons of differences in means with pre-test English vocabulary (form B only) ▶ Correlation analysis with other measures
Sentence comprehension test Norwegian vocabulary K-Bit 2 Riddles section K-Bit 2 Matrices section Memo game Background questionnaire	<ul style="list-style-type: none"> ▶ Between group comparisons of differences in means ▶ Correlation analysis with outcome measures 	Sentence comprehension test Sentence repetition test Delay condition yes/no decision task	<ul style="list-style-type: none"> ▶ Chronbach's Alpha ▶ Between group comparisons of differences in means ▶ Multiple regression for influence of independent variables ▶ Correlation analysis with other measures <ul style="list-style-type: none"> ▶ Chronbach's Alpha ▶ Between group comparisons of differences in means ▶ Correlation analysis with other measures, and between short and long sentences in the test ▶ Comparisons of differences in means for individual words in the test ▶ Comparisons of differences in means for long and short sentences <ul style="list-style-type: none"> ▶ Chronbach's Alpha

BACKGROUND MEASURES

A background questionnaire was administered to all parents in each student group asking for information about factors such as students' exposure to English-language media, time spent in an English-speaking country, number of siblings, and potential diagnoses which might influence L2 English acquisition. These data were used both to ensure that the two state-school groups were comparable on such factors, and for correlation analysis to check their influence on L2 English development.

Norwegian vocabulary was tested through a test produced specifically for the PhD study, based on form A of the Peabody Picture Vocabulary Test IV, which is described in more detail below. This test was administered in the pre-test session. Verbal and non-verbal ability was tested with the Matrices and the Riddles sections², respectively, of the Kaufman Brief Intelligence Test, Second Edition (K-Bit 2) (Kaufman & Kaufman, 2004). Visio-spatial working memory was tested using a memo game task where participants were asked to find pairs of pictures, where number of attempts and time used to complete the task were recorded. These measures were administered to check that groups were similar on these factors, and for use in correlation analysis in order to investigate their influence on English acquisition.

English vocabulary knowledge at the start of the school year was measured through the Peabody Picture Vocabulary Test IV, form B, which is further described below, and sentence comprehension through the English Listening Lesson Library Online game (Eillo, 2013a, 2013b) where pictures of pets and fruits, respectively, were described by a male native speaker of American English, and participants were asked to pick the picture matching this stimuli out of a set of three pictures. Since no norming exists for scores on this game, it was also administered to the control group.

THE PEOBODY PICTURE VOCABULARY TEST

The PPVT-IV is a standardized test of receptive vocabulary normed for native speakers of American English aged 2:6–90. Stimuli consist of individual words, and for each word, the participant selects the matching picture from a set of four. The number of words administered in the test depends on individual participants, as there is a discontinuation rule after a certain number of incorrect responses. The

2. The K-Bit 2 is a test of verbal and non-verbal intelligence. The Riddles section measures expressive (productive) vocabulary, while the Matrices section is a non-verbal pattern-recognition task.

test comes with detailed instructions for administration, as well as for points of reference for comparison with the reference group norms. Notably, raw scores on the test can be compared to both Growth Scale Values (GSV) and age equivalents for the reference group. Since such norming exists for the test, it was not administered to the control group in the PhD study.

THE SENTENCE COMPREHENSION TESTS

The post-test sentence comprehension test was designed specifically by the researcher and the supervisor for the PhD study. It was administered via a computer like the pre-test comprehension test, and consisted of 30 simple sentences read alternately by a female native speaker of British English and one of American English, both using standard varieties. For each sentence, the participant was asked to pick a picture that corresponded to the sentence out of a set of four pictures. Only one of these pictures corresponded completely to the sentence, while two others corresponded to parts of the sentence. The fourth picture was a distractor that did not correspond to any part of the stimulus. The test was designed to measure comprehension of vocabulary in context, and was also administered to the control group. Internal consistency was found to be good through Cronbach's alpha ($\alpha=.906$).

THE SENTENCE REPETITION TEST

The second test developed specifically for the study was a sentence repetition test. Its aim was to test grammatical processing, which proved difficult through a comprehension test with the language pair Norwegian/English, since both languages generally mark syntactic roles through (similar) word order. The repetition test consisted of 17 sentences of various lengths, which were read by the researcher and which each participant was asked to repeat one by one. The test session was audio-recorded, and performance was evaluated by two independent raters based on the number of morphemes correctly repeated. Pronunciation was not evaluated as long as it was comprehensible. The repetition test was also administered to the control group. Internal consistency was found to be good ($\alpha=.896$), as was inter-rater reliability as checked with Intraclass Correlation Coefficient (ICC =.953).

THE DELAY CONDITION YES/NO DECISION TASK

The delay condition yes/no decision task asked participants to respond to a total of 14 simple yes/no questions based on pictures that they saw three seconds after hearing the question. The intention was to force language processing by requiring participants to hold the question in working memory before responding. This test was also administered to the control group. However, internal consistency was low ($\alpha=.553$), and since other problems were also identified in this test, its results were not analyzed further.

DATA ANALYSIS

All three learner groups were compared through statistical analysis of variance (ANOVA) with planned comparison on post-test English vocabulary as measured by combined scores on both forms of the PPVT-IV. Scores on the sentence comprehension and sentence repetition tests were compared using the same methods for all four groups, including the control group. The two state-school groups' scores on vocabulary (PPVT-IV forms A and B combined), sentence comprehension, and sentence repetition were compared using t-tests, and for these two groups, multiple regression analysis was used to investigate the impact of group (i.e., volume of input) on sentence comprehension. For scores on the PPVT-IV form B in the two state-school groups, within-group comparison was performed using Wilcoxon signed ranks test, and between-groups comparison using Mann-Whitney U.

Results from the repetition test were analyzed to investigate whether it was a valid measure of language competence, and if so, what subskills it tested. ANOVA with pairwise comparisons compared the success with which participants repeated sentence initial, medial, and final words, respectively, since if repetition only entailed parroting, participants would be expected to repeat sentence final words more successfully. Furthermore, factorial analysis with group as the between-subjects factor and sentence length as the within-subjects factor looked for effects for sentence length, to see whether there was evidence that short sentences could be repeated as acoustic images without any language processing. In order to investigate which specific language ability was tested, ANOVA was performed on the two state-school groups' scores comparing words categorized into three groups. Functional words such as determiners and auxiliaries constituted the first group; the second group consisted of lexical words such as nouns, verbs, and adjectives which were overtly inflected, e.g., for number or tense, and the final category consisted of lexical words which were not overtly inflected, i.e. whose form was identical to their base form.

FINDINGS

On non-English background measures and pre-test English, statistical comparison found no significant differences between the two state-school groups. Thus, any differences in outcome measures between these groups were not likely to stem from differences in background. Between the two state-school groups on the one hand and the immersion group on the other, there was a significant difference for Norwegian vocabulary; pre-test English was not compared for the immersion group.

The overall results on post-test English measures are summed up in Table 17.2 showing the mean scores per group. They imply a relationship between input and performance in the post-test, as the immersion group outperforms the bilingually based group, which in turn outperforms the L1-based group on every measure of English competence in this study after one year of school. Results from the pre-test of vocabulary (PPVT-IV form B) are also provided as a point of reference, as this was the only test that was administered in both test sessions.

TABLE 17.2. Results for all groups on pre-test English vocabulary and post-test measures of English.

	L1-based group	Bilingually based group	Immersion group	Control group	Maximal possible score
PPVT-IV form B pre-test	23.7	25.4	47.4	–	–
PPVT-IV form B post-test	29.1	44.1	83.3	–	–
PPVT-IV form A post-test	33.9	40.2	77	–	–
Sentence comprehension	14.4	18.7	28.1	29.3	30
Sentence repetition	85.1	93.9	132.6	138.2	142

Note: For the PPVT-IV, age equivalents and Growth Scale Values as given in the test manual provide more useful points of reference than maximum scores, and these are discussed below.

The ANOVA for combined scores on both forms of the PPVT-IV found significant differences both between the immersion group and the bilingually based group, and between the bilingually based group and the L1-based group. When scores on

the sentence comprehension and sentence repetition tests were compared using the same methods for all four groups, including the control group, significant differences were found between the control group and the immersion group, between the immersion group and the bilingually based group, and between the bilingually based group and the L1-based group. An overall finding is thus that which group participants belonged to, and thus presumably volume of input, significantly impacted their performance in English after one year.

The two state-school groups were compared further, leaving out the immersion group since it was small and heterogeneous, and since this group's performance in English in the pre-test was notably higher than the two state-school groups. T-tests found significant differences between the two groups' performance on all English measures in the post-test sessions, i.e. vocabulary (PPVT-IV forms A and B combined), sentence comprehension, and sentence repetition. An important finding was that for the state-school groups, sentence comprehension showed the largest between-group difference. Multiple regression analysis was used to investigate the impact of English input in school inside and outside of English classes on this measure compared to the impact of pre-test scores on English vocabulary and on Norwegian vocabulary. Volume of input, i.e. which group students belonged to, was found to account for 24% of variance on the students' test scores, while their scores on the Norwegian vocabulary test and the pre-test for English vocabulary combined only accounted for 14% of variance.

Comparison of student scores on the English vocabulary test (PPVT-IV form B) using Mann-Whitney U confirmed that there was no significant difference between the two state-school groups on the pre-test, while on the post-test, there was. Repeated-measure Wilcoxon signed ranks tests within each of these groups found no significant difference in the L1-based group between performance on the English vocabulary pre-test and the post-test, indicating that no measurable development had taken place in the eight months between the two test sessions. The difference in the bilingually based group, on the other hand, was significant. As already mentioned, the PPVT-IV comes with a test manual with normed score reference points, making it possible to compare any individual or group score to the age at which this would be the average score in the test's native-speaker reference group. Such comparison confirmed that the difference in scores between the pre- and the post-test for the L1-based group was very small: The (non-significant) increase in the L1-based group in the course of eight months is equivalent to three months' vocabulary development as normed for English native speakers. Furthermore, differences in scores can be translated into Growth Scale Values (GSV) using the PPVT-IV test manual, where a difference in GSV of 8 is considered sig-

nificant. In the L1-based group, the GSV difference was only 5, again indicating non-significant difference between the pre- and the post-test. Development in the bilingually based group over the course of the same time span, however, was equivalent to a GSV of 12, and to 10 months' development in age equivalents as normed for the test manual for children who are native speakers of English. This means that the L2 English learners in the bilingually based group in fact developed more rapidly than what is typical for younger native speakers at the same level of vocabulary development in English.

For the immersion group, development on the English vocabulary test (PPVT-IV form B) was found to range from an age equivalent of 11 months to 37 months for individual participants in the course of the eight-month period between test sessions. Measured in GSVs, however, there was still a significant difference in the post-test between this group and the reference group for the test on both form A (GSV difference of 22 points) and form B (GSV difference of 21 points). Since a GSV of 8 is considered a significant difference, this clearly indicates that the immersion group had not reached native averages in English vocabulary comprehension.

Results on the sentence repetition test were analyzed for all groups in the discussion of whether this test constitutes a valid measure of language competence. With Pearson's r , scores on the test were found to be correlated with the other measures of English, and with Norwegian verbal ability, i.e., vocabulary and K-bit 2 Riddles. However, results on the sentence repetition test were not correlated with non-verbal measures, i.e., non-verbal ability as measured by K-Bit 2 Matrices, nor with working memory as measured by the memo game task.

The comparison between scores on initial and final words in the repetition test found no significant differences. Furthermore, the analysis of performance on short and long sentences indicated that performance on short sentences also differed between groups. Together, these findings indicate that participants were not merely parroting an acoustic image, and that the test was indeed a valid measure of language competence. The analysis of how well participants repeated functional, uninflected lexical, and inflected lexical words showed that the bilingually based group was significantly better at repeating lexical words with no overt inflection compared to lexical words with overt inflection and functional words, while in the L1-based group no such significant difference was found. The conclusion was the ability to perceive lexical words was the main advantage of the bilingually based group in this test.

DISCUSSION: CONTRIBUTIONS TO THE ENGLISH DIDACTICS FIELD

The contribution of the PhD study lies mainly in exploring the role of input in early-start L2 classrooms, and in new insights into how language acquisition takes place in such a context. Furthermore, the study contributed to developing methods for testing overall L2 proficiency in young, pre-literate learners at low proficiency levels, where receptive competence may be more advanced than productive skills. Finally, although the actual activities that took place in the classrooms in the three groups investigated were not controlled, the findings have some clear implications for teaching.

EMPIRICAL CONTRIBUTIONS

Empirically, the main contribution of the PhD study is found in the different results in the post-tests in the three groups of early English learners. The within-group analysis of vocabulary comprehension between the pre- and the post-test for the L1-based group demonstrated that in the setting of standard English classes in Norwegian grade 1, acquisition is not guaranteed, revealing a possible problem with such early-start programs. This was evident in the lack of vocabulary development in this group shown in the small, non-significant change in PPVT-IV form B scores between the pre- and the post-test.

In the bilingually based group, the modest increase in language exposure compared to a typical classroom had a significant effect on vocabulary development, which was found to be slightly faster than for competence-matched, younger native speakers of English. This finding is important since it may have consequences for the ambitions of early-start foreign language learning programs. Input in such settings will always be limited, but it seems that there can still be significant benefits from an early start if sufficiently substantial input is provided. The increased input appeared to have a particularly strong effect on sentence comprehension, which means that it is likely in turn to be beneficial for further acquisition from naturalistic exposure.

For the seven children in the immersion group, attending school for a year in a new language led to rapid acquisition, with scores on vocabulary, sentence comprehension and sentence repetition approaching native-speaker means indicating acquisition across the board. Although there was great variation between participants, we see that L2 immersion at such a young age can lead to very fast development. In this group, individual differences are particularly crucial, and important findings include how development on the PPVT-IV form B between the pre- and the post-test ranged from an age equivalent of 11 months to 37 months. There

was no clear pattern to what caused fast development; both the fastest-developing child and the child with the least development were bilingual upon starting school in English, and neither the child with the lowest nor the child with the highest English vocabulary scores in the pre-test showed the fastest development in terms of age-equivalence in months. Although the number of participants in this group is too small to draw firm conclusions, there is no indication here that previous language background is a strong predictor for development in a new language.

THEORETICAL CONTRIBUTIONS

The results of the PhD study are also relevant for theoretical questions about the nature of L2 acquisition. In particular, they support the assumption that foreign language learning is essentially a form of SLA. Thus, it depends on the same factors and follows the same trajectories as naturalistic SLA, input plays a crucial role, and a young AoA should in principle be beneficial. The findings are also relevant for the question of exactly what constitutes language competence in an L2. While pedagogical perspectives often focus on the four skills of listening, speaking, reading, and writing, these describe overt language behavior, not underlying competence. The PhD study is mainly relevant to what would typically be categorized as the listening skill, and emphasizes that this in itself is a multifaceted language ability.

One finding was the speed of vocabulary acquisition displayed by some of the young L2 learners compared to L1 acquisition. This rapid L2 vocabulary development compared to native speaker norms may indicate that L2 lexical acquisition is supported by L1 knowledge. Importantly, the PPVT-IV only tests receptive vocabulary, and vocabulary breadth rather than depth. Jiang (2004), for example, suggests that in early adult SLA, entries for new words initially consist of information about form (phonology and morphology) only, with a pointer to the L1 translation equivalent for meaning and syntax. It is possible that what had been acquired by the bilingually based group in the present study was such a pointer to the L1 lexicon, and this process may be beneficial for early comprehension of new vocabulary in an L2. The rapid lexical acquisition may also reflect that mechanisms that support L1 lexical acquisition in young children, such as fast mapping (Shintani, 2011), are still available to learners in this age group. In combination with support from the already well-developed L1 lexicon, such mechanisms may facilitate L2 receptive vocabulary in young learners in a way that allows very rapid development. Although the use of the L1 in the classroom was not studied here, it is possible that strategic L1 use in teaching may facilitate such processes.

However, the results also highlight how acquisition does not simply entail learning vocabulary items as translation equivalents from the L1, or as isolated items. The results of the sentence comprehension and the repetition test suggest that the main benefit of increased exposure is comprehension of words in context, not isolated lexical knowledge, and thus that processes such as phoneme discrimination, lexical retrieval, and building “good enough” representations for comprehension (cf. Carroll, 2004; Clahsen & Felser, 2006; VanPatten, 2012) are important competencies which constitute early steps in L2 development. The study indicates that these skills develop as a result of exposure.

The results of the PhD study also shed light on the exact role of input in SLA. For example, while it is common to assume a distinction between *input* and *intake*, studies have found that acquisition of various aspects of language can take place in the absence of attention to the input in both children and adults (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Schachter, 1998; Truscott, 1998). The children both in the bilingually based group and in the immersion group in the present study had acquired both receptive vocabulary and sentence processing and comprehension skills from naturalistic input, with no systematic effort on the part of the teachers to draw students’ attention to form in this input. Input may certainly have been simplified, but such adapted input is still naturalistic and resembles the child-directed speech that is common in L1 acquisition. There was little evidence in the PhD study to answer the question of whether also more detailed grammar competence can be achieved through naturalistic input without focus on form. However, the fact that some children in the immersion group scored within the native speaker range on the sentence repetition test, which required correct reproduction on functional morphemes, may indicate that also grammatical accuracy can in principle be acquired from naturalistic input in young L2 learners.

METHODOLOGICAL CONTRIBUTIONS

The PhD study contributed methodologically in developing tests that may be useful for investigating early stages of acquisition in very young, pre-literate learners. With the exception of the PPVT-IV, test materials in the study were developed specifically for the study. In particular, three methods were developed and tested, namely the sentence comprehension test, the sentence repetition test, and the delay-condition yes/no question task.

Although the ELLLO game (Elllo, 2013a, 2013b) used in the pre-test session probably succeeded in checking that the two main groups had comparable linguistic competence at the outset, there were problems both in difficulty level and in

the odds of guessing correctly in this game, with only three alternative responses per picture. Thus, for the post-test session, a sentence comprehension test was created specifically for the study. This test was found to distinguish well between competence levels, although there was some evidence of ceiling effects in the immersion group.

As reported above, the sentence repetition test also showed good internal consistency and distinguished between the competence levels of the participant groups. The detailed analyses of scores on different items in the test also supported the conclusion that it was a valid measure of language competence, and did not just test parroting abilities. The final test developed for the study was the yes/no question task with a delay condition. This test had low internal consistency, did not distinguish well between competence levels, and allowed some children in all groups to achieve very high scores either by chance or because the test was too easy. Thus, results of this test were not analyzed in the study and the conclusion is that it was not a successful measure of language competence.

IMPLICATIONS FOR TEACHING ENGLISH

In the study, the researcher did not carefully control methods used in the classrooms, and results therefore cannot be used to directly recommend teaching methods. Still, the study has implications for English teaching. First, it is clear that although English was abundant in Norwegian society, at the time of the study young children could not be expected to encounter enough English outside of school for any significant acquisition to take place. Secondly, attending English class from age 6 does not guarantee acquisition unless those classes provide sufficient input, and English classes as mandated by the current curriculum, still in place at the time of writing this chapter with only minor revisions, apparently do not necessarily provide such input. However, third, the study also demonstrates that input need not be unrealistically massive, and that rapid acquisition is possible if teachers focus on providing English input whenever possible both during class and throughout the school day within the existing curriculum.

Thus, those who teach English to young children in Norway should plan activities where they speak English to the students whenever possible. Activities successfully used in the bilingually based group in this study included conducting morning meetings and classroom management throughout the school day in English, with translations into Norwegian when necessary. Importantly, no effort was made to exclude the use of Norwegian, and as speculated above, it is possible that strategic translation and explanation in L1 may have facilitated L2 vocabulary

learning. There was also no specific effort to increase students' productive use of English, although increased production on their part may have been a natural result of the increased input. During English class, frequent activities providing English input included teachers reading aloud or talking about pictures or objects that they had brought into the classroom. An implication of the study is that these kinds of activities are well suited to young foreign language learners, and may be a better choice than some activities included in typical workbooks for first grade.

SUGGESTIONS FOR FUTURE RESEARCH

The results of this PhD study indicated that volume of exposure is a crucial factor in early L2 acquisition. The curriculum under which the study was conducted is still in place at the time of writing this chapter; it is currently being revised and the details of the new curriculum are still unknown, but it is clear that numbers of hours for English teaching will remain low. The challenge of providing enough input for acquisition to take place is thus still highly relevant. The international debate about the interplay between AoA and input in foreign language learning is still ongoing (cf. DeKeyser, 2017), and further studies in the Norwegian context are needed.

Future research would benefit from including more participant groups and different teachers. A larger immersion group would also be beneficial in order to obtain generalizable knowledge about this setting. Future similar studies may also control more closely methods and activities in the different classrooms in order to gain a more detailed picture of exactly what practices lead to successful acquisition, and longitudinal studies may provide knowledge of the long-term effects of the different conditions. The increased input in English in the bilingually based group in this study was maintained through grade 2, and two master studies (Sivertzen, 2013; Strand, 2014) followed up the findings in the years after data collection in the PhD study, finding no advantage of the bilingually based group on the PPVT-IV in grade 4 compared to students who had received regular English teaching from grade 1, but still finding an advantage for sentence comprehension in grade 5.

The study highlights the importance of English Grade 1 teachers having the English proficiency and confidence to use English extensively in the classroom, and it is a step in the right direction that requirements have been introduced since this study was concluded for English teachers at primary level to have formal qualifications in English. However, as an added source of L2 exposure, further research may look into whether input from film and similar media can have simi-

lar beneficial effects for early L2 learners in classroom settings. Interestingly, a recent survey found that young Norwegian children watch English-language media at very young ages (TNS Kantar, 2017), and investigating whether such exposure can be used to promote acquisition in an instructed setting for the youngest learners might be useful. Finding ways to ensure that the early start with English in Norwegian schools is effective is still important.

REFERENCES

- Bialystok, E., & Hakuta, K. (1999). Confounded age: linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* (pp. 161–181). Mahwah, NJ: Lawrence Erlbaum Associates.
- Burston, C. (1975). French in the primary school: The British experiment. *Canadian Modern Language Review*, 31, 388–402.
- Carroll, S. E. (2004). Segmentation: Learning how to “hear words” in the L2 speech stream. *Transactions of the Philological Society*, 102(2), 227–254. doi: [10.1111/j.0079-1636.2004.00136.x](https://doi.org/10.1111/j.0079-1636.2004.00136.x)
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. doi: [10.1017/S0142716406060024](https://doi.org/10.1017/S0142716406060024)
- Dahl, A. (2014). *Young language learners: The acquisition of English in Norwegian first-grade classrooms*. (Doctoral thesis). NTNU, Trondheim.
- Dahl, A. (2015). Input and language competence in early-start foreign language classrooms. In T. Cadierno & S. W. Eskildsen (Eds.), *Proceedings: Thinking, Doing, Learning: Usage Based Perspectives on Second Language Learning* (pp. 125–151). Berlin: Mouton.
- Dahl, A., & Vulchanova, M. (2014). Naturalistic acquisition in an early language classroom. *Frontiers in Psychology: Special issue. Learning a non-native language in a naturalistic environment: Insights from behavioural and neuroimaging research.*, 5, 1–9.
- de Bot, K. (2006). The plastic bilingual brain: Synaptic pruning or growth? Commentary on Green, et al. *Language Learning*, 56(S1), 127–132. doi: [10.1111/j.1467-9922.2006.00358.x](https://doi.org/10.1111/j.1467-9922.2006.00358.x)
- DeKeyser, R. (2017). Input is not a panacea. 0(0), 1367006918768371. doi: [10.1177/1367006918768371](https://doi.org/10.1177/1367006918768371)
- Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test. Minneapolis, MN: Pearson Education, Inc.
- Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164–194. doi: [10.1093/applin/aml015](https://doi.org/10.1093/applin/aml015)
- Ello. (2013a). Listening Games: Fruit. Retrieved from <http://www.ello.org/english/Games/G008-Fruit.html>
- Ello. (2013b). Listening Games: Pets. Retrieved from <http://www.ello.org/english/Games/G039-Pets.html>

- García Mayo, M. d. P., & García Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language*. Clevedon: Multilingual Matters.
- Holmstrand, L. S. E. (1982). *English in the elementary school: Theoretical and empirical aspects of the early teaching of English as a foreign language*. Uppsala: Acta Universitatis Upsaliensis.
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturation constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–588). Oxford: Blackwell.
- Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, 88(3), 416–432. doi: [10.1111/j.0026-7902.2004.00238.x](https://doi.org/10.1111/j.0026-7902.2004.00238.x)
- Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman Brief Intelligence Test. Minneapolis, MN: Pearson Education, Inc.
- Krashen, S. D., Long, M. H., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, 13(4), 573–582.
- Krashen, S. D., Long, M. H., & Scarcella, R. C. (1982). *Child-adult differences in second language acquisition*. Rowley, MA: Newbury House.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: Wiley.
- Muñoz, C. (Ed.) (2006). *Age and the rate of foreign language learning* (Vol. 19). Clevedon: Multilingual Matters.
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44(01), 1–35. doi: [10.1017/S0261444810000327](https://doi.org/10.1017/S0261444810000327)
- Murphy, V. A. (2010). The relationship between age of learning and type of linguistic exposure in children learning a second language. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 158–178). London/New York, NY: Continuum International Publishing Group.
- Oller, J. W., & Nagato, N. (1974). The long-term effects of FLES: An experiment. *The Modern Language Journal*, 58(1–2), 15–19.
- Penfield, W., & Roberts, L. (1959). *Speech and brain-mechanisms*. Princeton, NJ: Princeton University Press.
- Piske, T., & Young-Scholten, M. (2009). Introduction. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 1–28). Bristol: Multilingual Matters.
- Pulvermuller, F. F., & Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Language Learning*, 44(4), 681–734. doi: [10.1111/j.1467-1770.1994.tb00635.x](https://doi.org/10.1111/j.1467-1770.1994.tb00635.x)
- Rosansky, E. J. (1975). The Critical Period for the acquisition of language: Some cognitive developmental considerations. *Working Papers on Bilingualism*(6), 1–12.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101–105. doi: [10.1111/j.1467-9280.1997.tb00690.x](https://doi.org/10.1111/j.1467-9280.1997.tb00690.x)
- Schachter, J. (1998). Recent research in language learning studies: Promises and problems. *Language Learning*, 48(4), 557–583. doi: [10.1111/0023-8333.00059](https://doi.org/10.1111/0023-8333.00059)
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15(2), 137–158. doi: [10.1177/1362168810388692](https://doi.org/10.1177/1362168810388692)

- Sivertzen, T. G. (2013). *The role of input in early second language acquisition: A study of the long-term effects of initial extra English input on Norwegian 4th graders' receptive vocabulary* (MA thesis), NTNU, Trondheim.
- Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and speech*, 20(4), 357–365. doi: 10.1177/002383097702000407
- Strand, O. L. (2014). *Do we have to speak English? A longitudinal perspective on the impact of early English input on the development of English comprehension skills in Norwegian pupils*. (MA thesis), NTNU, Trondheim.
- TNS Kantar. (2017). *Mediebarn 2017: En kartlegging av norske barns medievaner*. Retrieved from <https://www.sprakradet.no/globalassets/vi-og-vart/rapporter-og-undersokelser/mediebarn-2017-rapport.pdf>
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14(2), 103–135. doi: 10.1191/026765898674803209
- Utdanningsdirektoratet. (2006). Knowledge Promotion: Curriculum in English. In: Utdanningsdirektoratet, www.udir.no.
- VanPatten, B. (2012). Input processing. In A. Mackey & S. M. Gass (Eds.), *The Routledge handbook of second language acquisition* (pp. 268–281). London: Routledge.