Doctoral (PhD) Dissertation

The Long-Term Impact of Learner-Learner Interaction on L2 English Development

DOI:10.18136/PE.2021.776

Written by
Feisal Aziez

Supervisor
Prof. Dr. Marjolijn Verspoor

Multilingualism Doctoral School

Faculty of Modern Philology and Social Sciences

University of Pannonia

Veszprém, 2021

The Long-Term Impact of Learner-Learner Interaction on L2 English Development

Thesis for obtaining a PhD degree in the Doctoral School of Multilingualism of the University of Pannonia

in the branch of Applied Linguistics

Written by Feisal Aziez

Supervisor: Prof. Dr. Marjolijn Verspoor

Propose acceptance (yes / no)

…….
(supervisor)

As reviewer, I propose acceptance of the thesis:

Name of Reviewer: …..... yes / no ……
(reviewer)

Name of Reviewer: …..... yes / no ……
(reviewer)

The PhD-candidate has achieved ….....% at the public discussion.

Veszprém, ……. 2021 ……
(Chairman of the Committee)

The grade of the PhD Diploma …..... (……. %)

Veszprém, ……. 2021 ……
(Chairman of UDHC)

ABSTRACT

This dissertation aimed to investigate how extensive peer-to-peer interaction in a pesantren affects the learners' L2 development over time in one academic year. There are two cohorts involved in the study, a first-year group with 82 learners and a second-year group with 56 learners. This cross-sectional, longitudinal design was meant to simulate a two-year developmental path. Taking a dynamic usage-based (DUB) perspective of language learning, which holds that frequency of exposure and use is the strongest predictor in L2 development, we assumed that with so little authentic input and so much repetition of learners' non-target utterances that the learners might create their own version of English, which would eventually stabilize and be considered a pidginized version. Four interrelated studies were devised to explore pesantren learners' practices and language development.

The first study examined the learners' peer interaction, particularly in terms of interactional features which reportedly promote L2 acquisition including turn taking, trigger, negative feedback and modified output. Samples of learners' interaction were examined for these interactional features. The findings clearly indicate that peer interaction among the learners in the pesantren lacks the interactional features that can promote language learning.

The second study examined the effect of individual differences such as gender, age of acquisition, motivation and scholastic aptitude on the learners' L2 writing development. A LHQ, learners' reflection on motivation, and academic reports were used for this purpose. Gains were operationalized as the difference between beginning and end scores. A regression analysis shows that in Group 1, initial writing proficiency and age of acquisition were significant predictors of gains. Age of acquisition contributed negatively to the gains, which means the earlier they started learning English, the more gains. In Group 2 only the initial writing proficiency was found as a significant positive predictor. Gender and motivation, on the other hand, were not found to be strong predictors in either group. Scholastic aptitude did show a significant effect on gains in Group 1, but not in Group 2 when initial writing proficiency (covariate) was controlled for. However, scholastic aptitude was significant when the covariate was excluded.

The third study explored English development of learners over time with bi-weekly writing. The statistical analyses showed that Group 1 improved significantly in the first half of the year and then stabilized. Group 2 was significantly better than Group

1 only in the first scores at the beginning of the academic year. The first group showed significant improvement in the first semester but not in the second semester. In Group 2 there was no significant difference between pre, mid and post scores. This means that the learners in Group 2 did not make any significant progress during the one-year period. A further regression analyses was performed with gains as the outcome variable and variability, class ranking and initial proficiency as predictors. Results show that variability was a significant predictor of performance on the writing test in both Group 1 and Group 2.

In the fourth study, the aim was to explore the extent of fossilization or pidginization in the learners' L2 in the context of pesantren. Sample texts were examined for the characteristics of pidginization. The findings show strong indications of pidginization in the learners L2 starting after the first semester in the first year. Learners in Group 1 show that at the beginning they have many more Pidginization forms (P-forms), than they do later on as they improved significantly by producing a lower pidginization ratio overtime. However, the longitudinal analysis shows that the substantial improvement occurred mostly in the first few sessions only and then seem to stabilize. We also counted types of pidginization features and found that the groups produced a rather similar percentage in each feature.

Together the findings suggest that learners make almost all progress in the first six months and then they stabilize in the forms and expressions that they use, which may be considered a fossilized system with typical pidginization features. Apparently, as the learners feel that they have a repertoire sufficient to communicate with each other, they do not make much progress anymore. During their interaction the NTL output they produced was rarely corrected, probably because the learners had no clue that the forms were not target-like. It was also clear that the learners in the pesantren have only limited exposure to authentic or expert L2 input as the input they receive is mainly from their peers. Moreover, the type of instruction they receive from their teachers is mainly lexically based. These factors may cause the learners progress to stagnate, as the developmental part of this study suggested. Finally, the findings of Study 4 also suggest a role for the extensive peer interaction in promoting pidginization process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
BLCLAB	Brain, Language, and Computation Laboratory
BT	Baby Talk
CAF	Complexity, Accuracy, Fluency
CDST	Complex Dynamic System Theory
CLIL	Content Language Integrated Learning
CoV	Coefficient of Variation
DUB	Dynamic Usage Based
EFL	English as a Foreign Language
ESL	English as a Second Language
FFE	Focus-on-Form Episodes
FT	Foreigner Talk
FTF	Face-to-Face
FUMM	Form-Use-Meaning Mappings
HL	Heritage Language
IL	Interlanguage
ISLA	Instructed Second Language Acquisition
L1	First Language
L2	Second Language
LHQ	Language History Questionnaire
LRE	Language-related episodes
NNS	Non-Native Speaker
NS	Native Speaker
NTL	Non-target
P-Forms	Pidginization Forms and Constructions
SCMC	Synchronous Computer-Mediated Communication
SDT	Self-Determination Theory
SLA	Second Language Acquisition
SLD	Second Language Development
SPSS	Statistical Package for the Social Sciences
TL	Target-like
UBL	Usage-Based Linguistics
WTC	Willingness to Communicate

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Marjolijn Verspoor. Her guidance, motivation, patience, and immense knowledge have helped me since the first day of my PhD study. This dissertation would not be a reality without her amazing support. Her enthusiasm always lifted me up when I was lost during the research and writing process. I could not imagine having a better supervisor.

I would also like to extend my profound appreciation to everyone at the doctoral school who have guided me through my PhD journey and formed me as a learner in the field of Applied Linguistic in general and specifically in Multilingualism. They have set not only professional, but also personal examples for me to follow in the future. Their kind support and patience have inspired me from the first day of my PhD journey.

I would also like to express my gratitude to the principal of the pesantren for allowing me to conduct my research there, to the English teachers at the pesantren who assisted me in collecting the data despite all the problems, and to all the amazing students at the pesantren for their willingness to be involved in the study.

I would like to thank the Hungarian Government for providing financial support through the Stipendium Hungaricum program so that I could pursue my PhD in Hungary.

Finally, I am also grateful to all my fellow Indonesian students, my fellow PhD students at the doctoral school, and my fellow international students in University of Pannonia for the companionship during my life in Hungary. There are also many others who I cannot mention one by one who have supported me in the journey. Without all these people, completing this research would have been impossible.

CHAPTER 1
BACKGROUND LITERATURE

This dissertation will explore the English language development of 138 young Indonesian learners in their first and second year at a *pesantren*, an Indonesian Islamic boarding school, which promotes English learning especially through peer interaction. If we consider language development from a usage based theoretical perspective, frequency of exposure and experience are the main drivers of language development. The learners at the pesantren have little access to authentic English and the danger may be that they rely too much on their own interactions for input and output without authentic examples, which may lead to fossilization and pidginization. This chapter presents the background literature, the context and the theoretical positions of this dissertation.

1.1. Introduction

Peer interaction or learner-to-learner interaction has been widely used in second or foreign language classrooms across the globe to facilitate learners in order to improve fluency in the target language. In most cases, peer interaction is implemented through common classroom practices such as drills or information gap exercises. Several studies have supported the practice by indicating that peer interaction can promote L2 acquisition particularly in a psychological sense where learners feel less anxious in expressing their thoughts in L2 in comparison to learner-teacher interaction (e.g., Philp, Adams, & Iwashita, 2014; Loewen & Sato, 2018; Philp et al., 2014). However, most of these studies are conducted in laboratory or classroom settings in which the interaction is manipulated in some ways by the researchers and carried out in a relatively short period of time (e.g., Mackey 2012; Loewen 2015). In a meta-analysis, Mackey and Goo (2007) found that from 28 studies that they analysed, 64% of them were conducted in laboratory settings, while the rest were conducted in classroom settings. Additionally, these studies generally examined the features of interactions during negotiation of meaning and how they affect L2 learning (Loewen & Sato, 2018). So far, however, there has been little discussion on the long-term impact of peer interaction, especially of that taking place in naturalistic settings. This is because it is sometimes difficult for researchers to manage the complexities of the variables in the naturalistic classroom context (Shadish, Cook & Campbell 2002).

However, in Indonesia, there is a relatively unknown educational system called pesantren, which may allow researchers to investigate the impact of extensive L2 peer-interaction on the learners in the long run. This is made possible because students in a pesantren live and study within a school complex. Moreover, some pesantren institutions in Indonesia require their learners to communicate in the target languages (i.e., English and Arabic) outside the classrooms. It should be noted that although not every pesantren institution in Indonesia obliges their students to use L2 in daily communication, such practice is widely found across the country especially in the pesantrens that have adopted a modern curriculum (see Bin Tahir, 2015, 2016; Bin Tahir, Atmowardoyo, Dollah & Rinantanti, 2017; Jubaidah, 2015; Aziez, 2016; Risdianto, 2016; Raswan, 2017). A further discussion of differences among pesantrens is beyond the scope of the current study. However, peer-interaction in the L2 in the context of a pesantren is different from the practice at any other educational institution. Not only is it used as a form of language learning, but also as a form of daily communication to exchange meaning. Moreover, as is clear from observation, learners in a pesantren spend significantly more time communicating with their peers than with their teachers, who are more proficient L2 speakers. Thus, the majority of the learners' input is received from their peers and not from authentic or more proficient sources.

These conditions raise some questions on how the learners' L2 develops with such extensive peer-interaction. In recent theories on language development, it has been argued that authentic exposure as well as frequency are important in the success of language acquisition. For instance, in a dynamic usage based (DUB) approach (see Verspoor & Behrens, 2011: 38), the target language is seen as a set of conventions and learners will pick up the conventions that they hear most frequently. Therefore, it is important to give learners as much authentic input as possible. However, in a pesantren, learners tend to get their input from their peers and may pick up the conventions that they hear most frequently from each other. In a previous descriptive study describing the learners' English in a pesantren (Aziez, 2016), the learners' English contains a preponderance of L1 interference forms and overgeneralizations at the lexical, syntactical and phonological levels.

As mentioned earlier, peer-interaction has been argued to support language learning to some extent, but it is not without criticism. Some researchers believe that corrective feedback from peers can be poorer in quality compared to feedback from the teachers (Adams, 2007). Xu, Fan, and Xu (2019) also reported that learners tend to be

more hesitant in providing corrective feedback to their peers. They also found that the learners provided more corrective feedback on morphosyntactic errors than lexical and phonological errors.

The aforementioned studies as well as the description of the pesantren lead to the question whether the language that the learners in a pesantren produce becomes fossilized and may be considered a pidginized form of English. According to Richards (1974: 77), there are similarities between learners' languages and pidgin languages. Both codes are seen as a result of language contact and characterized by grammatical structure and lexical content originating from two or more languages. This notion led Schumann (1978) to his study on Alberto, a Spanish speaking immigrant in the US. In his study, which gave birth to the acculturation hypothesis or the pidginization hypothesis, he concluded that a pidginized form of a language may develop for two main reasons; (a) when learners separate themselves socially and psychologically from speakers of the target language, and (b) when the target language is used by learners for a very limited range of functions (Richards & Schmidt, 2010). In a later study, Andersen (1981) compared Alberto's English IL and Bickerton's (1977) research on Hawaiian Pidgin English and found similarities between both types of linguistic codes.

Since pidgin languages are used primarily for communicating ideas, they are restricted languages that serve only a communicative function; speakers of pidgins normally do not identify themselves with the group who speak the pidgin. They tend to reside in their own group apart from purposes of contact with the other group. This is not really the same in the case of learners in a pesantren. Since they are forced to speak English inside the school complex, English is used primarily to communicate ideas and they do not identify themselves as English speakers but they do form a speech community and group within the pesantren. This is similar to the case of a pidgin-like language produced by students in immersion programmes in Canada and the United States (Swain, 1997; Hammerly, 1991). Being critical of this type of communicative approach, Hammerly (1991) especially scrutinized these immersion programmes and concluded that although the students were successful in attaining a high level of communicative proficiency (fluency), they failed in terms of linguistic accuracy. He cites studies which show that "an error-laden classroom pidgin becomes established as early as Grade 2 or 3 because students are under pressure to communicate and are encouraged to do so regardless of grammar" (1991: 5).

On that basis, the present study aims to examine the development of English learners in a pesantren, which relies heavily on peer-interaction in the learning process without much authentic exposure. This study will also seek whether this condition will result in stagnation in their L2 development and exhibits features of pidginization.

Section 1.2 describes in detail an education system in Indonesia named pesantren and brings an overview of language learning practice in pesantren institutions in Indonesia. Section 1.3 deals with the role of interaction and second language acquisition, consisting of the general theories and previous studies from interactionist approach. Section 1.4 provides a discussion on second language development from a dynamic usage-based perspective. Section 1.5 deals with second language acquisition and the issue of pidginization, emphasizing the comparison between the two concepts. Section 1.6 concludes this chapter by summarizing the relevant theoretical positions and presenting the questions that the current study aims to answer.

1.2. Language learning in a pesantren

As mentioned previously, the unique context of a modern pesantren in Indonesia could provide an opportunity to see the extent to which extensive practice of peer interaction affects L2 development. Therefore, it is important to first understand what is a pesantren and why the current study focuses on this particular context. According to an Indonesian encyclopaedia on education, the term *pesantren* or *pondok pesantren* means a gathering place to learn Islamic teaching (Poerbakawaba, 1976). The term is commonly translated into English as Islamic boarding school. Ziemek (1986) believed that the term pesantren comes from its root word *santri* which mean pupil. In a pesantren, the pupils come and learn from the teachers whom they address as *kiai* or *ustaz* (Ahmad, 2012). The Pesantren is one of the Indonesia' oldest religious learning traditions and its existence can be traced back to the fifteenth century (Umar, 2014). At that time, the pesantren was the only educational institution helping society in improving literacy (Qomar, 2005). It is considered as the foundation of the indigenous educational system of Indonesia. Besides its huge base on Java Island, pesantren institutions are spread also on the outer islands of Indonesia as well as the Malay Peninsula (van Bruinessen, 1994). Its numbers are growing continuously. According to the Indonesian Ministry of Religious Affairs (2020), there are more than 27,000 institutions in the country, around 82% of which on Java Island, accommodating more than 4 million students.

In contrast to other educational institutions in the country, students in a pesantren typically live and learn inside or near the institutions with the teachers (Hidayat, 2007; Daulay, 2009; Bin Tahir, 2015, 2016; Bin Tahir et al., 2017; Jubaidah, 2015; Aziez, 2016; Risdianto, 2016; Raswan, 2017). Furthermore, while most schools in Indonesia are under the regulations of the Ministry of Education and Culture, these schools operate under the Ministry of Religious Affairs. According to (Dhofier, 1985), generally, there are two different types of pesantren. The first type is the traditional pesantren (also called *salafi*), which teach Islamic religion exclusively. The second type is the modern pesantren (also called *khalafi*), which in the past few decades has begun adopting a contemporary education system—teaching the students common subjects including English (Zakaria, 2010). The modernization of the institution is also reflected in the use of technology in its educational practices (Wekke & Hamid, 2013). As mentioned earlier, a detailed discussion on the different types of pesantren is beyond the scope of this study and we will focus on one particular type of pesantren.

In many modern pesantrens, there are usually three languages used as medium of instruction in the classrooms: Bahasa Indonesia, Arabic, and English (Bin Tahir, 2015). Indonesian is used in subjects included in the national curriculum such as mathematics, physics, chemistry, social science, civic education, etc. Arabic is used mainly in Islamic subjects and Arabic language subjects such as *nahwu* (syntax), *sharaf* (morphology), *fiqh* (Islamic jurisprudence), *tafsir* (commentary on the holy book), *muthalaah* (learning to learn), etc. While these two languages are used extensively in a large number of subjects, English is taught only in English language related subjects. Although some schools have adopted foreign languages other than Arabic and English (e.g., German, French, or Japanese), these two foreign languages still receive the most attention from modern pesantren institutions in their curricula because of the significance of both languages. Arabic is the language of the *Quran* and *Hadits*, the primary source of Islamic teachings, and therefore it is very important for the students to learn Arabic in order to better understand them. English, on the other hand, is perceived as the language of science and global communication. Moreover, a study by Farid and Lamb (2020) revealed that learning English also has a spiritual motive for the students in a pesantren, i.e., to be able to use English as a tool of da'wah (Islamic propagation) and to be able to communicate with other Muslims worldwide.

What is unique about this system compared to conventional schools is the extent of the use of these foreign languages. In addition to the use of Arabic and English as the

mediums of instructions, many modern pesantrens in Indonesia oblige their students to use English and Arabic, interchangeably on a weekly basis, in their daily communication inside the school complex. Since they study and live there, it means that they have to speak either English or Arabic at all times during the respective weeks (see Bin Tahir, 2015, 2016; Bin Tahir et al., 2017; Jubaidah, 2015; Aziez, 2016; Risdianto, 2016; Raswan, 2017). One of the pioneer pesantrens that obliged their students to speak English and Arabic instead of Indonesian and the local language is the pesantren of Gontor in East Java Indonesia (van Bruniessen, 2006). For decades, its graduates have spread and become teachers in pesantren institutions across the country and applied the same policy. Indonesian and local languages are usually allowed to be used in daily communication only in the first few months after the students' enrolment in the school. After that, both languages are strictly limited—allowed only in classes in which the language of instruction is Indonesian or the local language and when they talk to people who work in the school except the teachers. 'Illegal' use of Indonesian or local language by the students will lead to punishment. The forms of punishment given to the students vary. In the past decade, for example, it was common to hit, with a rattan stick, those students who break the school rules, the frequency of which depends on the severity of the violation. However, such practices have been disappearing from pesantrens. They are now moving towards more 'educational' punishments where, for example, students are asked to memorize 60 words in Arabic or English and their meanings in Indonesian (e.g., Jihad, 2011). Students who have been punished are then assigned to be *jasus* (literally translated as spy) who have to lookout if any of their friends speak Indonesian or the local language.

Although in most pesantren institutions there are two foreign languages being taught, this dissertation will focus only on English. As described earlier, the teaching of English in a pesantren is different from that in other school systems in Indonesia. In most conventional schools, English is taught and practiced only in the classrooms. English teachers in Indonesia struggle to accommodate their students in English classes because of limited instruction time, especially after the implementation of the 2013 National Curriculum in which time allotment for English as a subject was reduced (Panggabean, 2015). Although both systems follow the same curriculum, pesantrens also have their own curriculum focusing on language and religious subjects (Sofwan & Habibi, 2016). For instance, the National Curriculum allocated only two lesson hours (80 minutes) for English class every week. However, in many pesantren institutions, the students get another additional two lesson hours (80 minutes) of English reading class, which is part

of the school curriculum. Moreover, since the students in pesantrens live inside the institutions, schools have more flexibility in developing their own extracurricular activities. This allows for more input in learning English and more chance for them to practice their English.

There have been several studies exploring the practice of language learning in pesantren institutions in Indonesia. In a descriptive study, Bin Tahir (2016) explored the approaches of multilingual teaching and learning methods used in three pesantren institutions in Makassar, Indonesia. Based on his observation, all institutions in the study implemented a combination of an immersion approach, where the learners were taught in the target languages (i.e., English and Arabic) from day one, especially in the subjects that belong to the pesantren curriculum. He noted four main strategies used by the institutions to promote language learning. The first strategy is through teacher-student communication, where the teachers are engaged in the learning activities, which generally occur in the classrooms. The next strategy is the practice of learner-learner interaction both inside and outside the classrooms which, as Bin Tahir described, occurred "without error correction by the teacher or other students" (2016: 90). The institutions also applied a language specific rule where learners had to communicate in the target language(s) in their daily routines. Finally, several group activities were also implemented by the institutions including *muhadharah* (public speaking practice), language camps, and language clubs.

Another study by Al-Baekani and Pahlevi (2018) reported similar practices in one pesantren in West Java, Indonesia. They observed that the pesantren applied a Community Language Learning model, which emphasizes a communal sense in the learning group and encourages interaction as a means of language learning. However, the language learning model in the pesantren was not developed based on a syllabus or textbook but was transferred from generation to generation. The language teachers even claimed that they were not aware of any kind of model applied at the pesantren, which is also the case in Bin Tahir's (2016) study. The teachers developed the learning materials based on their own life in the pesantren and relied on learners' conversations in their daily activities to entrench the target language(s).

Indeed, studies on the language learning practices in a pesantren have only been carried out recently despite the fact that such practice in pesantren institutions is common in the country and has been around for decades. This is due to the fact that most such research has focused on language learning in conventional educational systems and little

attention has been paid to religious educational institutions such as the pesantren. Recent studies have documented English and Arabic language learning in different islands across the country including Java (e.g., Hidayat, 2007; Aziez, 2016; Al-Baekani & Pahlevi, 2018), Sumatra (e.g., Ritonga, Ananda, Lanin & Hasan, 2019), Sulawesi (e.g., Bin Tahir 2016; Bin Tahir et al., 2017), and even in Papua (e.g., Wekke, 2015) where Muslims are the minority. One point that has been consistently reported by these studies is the emphasis on peer-interaction in the language learning practice in pesantren institutions. In a previous study by Aziez (2016), such practice has been reported to result in non-target-like L2 production by the learners. However, how the learners in a pesantren interact and the extent to which the learners' develop in their L2 have not been well-documented.

The above description of the pesantren provides only a general picture of what pesantren institutions are and what language learning practices take place in the institutions. A more detailed description of the pesantren institution where the current study was conducted will be provided later in the next chapter.

1.3. Interaction in second language acquisition

For the past few decades, a lot of research has been carried out to understand the role interaction plays in second language acquisition (SLA). However, the importance of interaction in SLA had been overlooked before the introduction of the interaction hypothesis first articulated by Long (1981, 1983), which he revised later in 1996 (Long, 1996). Long basically stated that conversational modifications (i.e., comprehensible input) in an interaction between two or more people can promote acquisition. It is argued that when L2 learners engage in an interaction and face communication problems, they have the opportunity to negotiate solutions, which therefore facilitate acquisition of the target language. Although this construct has been largely credited to Long, it was principally based on discourse analysis studies during the 1970s (e.g., Wagner-Gough & Hatch, 1975; Hatch, 1978).

Another relevant theory emphasizing the need for comprehensible input in SLA was the theory from Krashen (1982), suggesting that comprehension of message meaning is important for L2 learners in order to internalize target language forms and structures. Krashen coined this notion as the "input hypothesis", which is constructed on both input and interactional modifications. Both Long and Krashen highlight comprehensible input as a source of acquisition. Although Swain (1985) recognizes the importance of

comprehensible input, she argues that it is not sufficient. She, therefore, developed what is called "comprehensible output" also known as the "output hypothesis", which suggests three functions of leaners' output, which focuses on accuracy rather than fluency. The first function namely the noticing function is elaborated by Swain (1995):

> In producing the target language (vocally or subvocally) learners may notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially, about the target language. In other words, under some circumstances, the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2. (p. 125-126)

The second function is called the hypothesis-testing function. When a learner says something in the L2, there is always an implicit hypothesis in his or her utterance, e.g., about the grammatical form of his or her utterance. By expressing himself or herself through that utterance, the learner tests this hypothesis. When he or she receives feedback from an interlocutor, the learner may reprocess his or her hypothesis. The metalinguistic function, the third function, is a conscious reflection by learners on the language they learn when they produce L2 utterances, which enables them to control and internalize linguistic knowledge.

Since Long first proposed the hypothesis, it has evolved into a theoretical approach (Mackey & Gass, 2015), which includes a description of multiple processes related to L2 learning (Mackey, 2012; Pica, 2013). These processes include exposure to the target language (input) and production of the target language (output) and their interaction with learners' cognitive resources and other individual differences (Long, 1996; Gass, 1997; Mackey, 2012; Pica, 2013; Gass & Mackey 2015; Long, 2015; Loewen & Sato, 2017). The earlier interactionist studies focused on how interaction is carried out in different settings. Some of the topics including speech modifications and interactions between native/non-native speakers as well as non-native/non-native speakers (Gass & Varonis, 1985; Varonis & Gass, 1985; Doughty & Pica, 1986; Porter, 1986; Pica, 1988; Gass & Varonis, 1990; Loschky, 1994). Researchers were particularly interested in how the interactants negotiate meaning—the frequency, the influencing factors, and its process (e.g., Long & Porter, 1985; Pica et al., 1991; Pica, 1994; Lyster & Ranta, 1997).

These studies have helped to reveal the characteristics of interaction, which consequently allow researchers to investigate specific variables related to interaction. Some of the most notable interactionist research studies, for example, focus on (a) discourse moves e.g., modification of input (Swain, 1985, 1995, 2005), (b) cognitive constructs e.g., noticing (Schmidt, 1990, 1995, 2001), and (c) L2 development and acquisition (Mackey, 1999; Spada & Lightbown, 2009; Mackey, 2012). The investigated variables are generally categorized into four domains: those concerning (a) the interlocutors (e.g., L2 proficiency, L1 status, gender, etc.), (b) the task characteristics (e.g., complexity, type of task), linguistic targets, and (d) the interactional context (Loewen & Sato, 2018). Since then, many researchers have moved their focus from investigating the general effectiveness of interaction to exploring the effectiveness of specific components of interaction in relation to the context and L2 learners.

The interest in interaction has been growing since its first emergence with numerous subsequent empirical studies in the forms of reviews (Gass, 2003; Plonsky & Gass, 2011; Goo & Mackey, 2013; Lyster & Ranta, 2013; Lyster, Saito & Sato, 2013; Plonsky & Brown, 2015; Kim, 2017) and meta-analyses (Russell & Spada, 2006; Mackey & Goo, 2007; Li, 2010; Lyster & Saito, 2010; Brown, 2016; Ziegler, 2016) investigating both general and specific components of interaction. These studies generally indicated the benefits of interaction for L2 acquisition. For instance, a meta-analysis of 14 quasi-experimental studies on interaction by Keck et al. (2006) have discovered a significant positive effect of interaction on L2 learners in the immediate posttests. Another meta-analysis of 28 interaction studies conducted inside and outside the classroom settings by Mackey and Goo (2007) also indicated a positive effect of interaction on L2 learning. This effect is reported to be more apparent on delayed posttests. In order to better understand about the concept of interaction in L2 acquisition, the key components of interaction will be presented below.

1.3.1. Components of interactions

1.3.1.1. Input

In the interactionist approach, input is a vital component of acquisition from which learners can derive linguistic hypotheses (Gass & Mackey, 2020). Gass and Mackey, (2020) defined it simply as the exposure to target language in a communicative context. Interactionist researchers have been particularly interested in the kinds of input received by L2 leaners namely naturalistic, pre-modified, and interactionally modified

input (Loewen & Sato, 2018). The main reason behind modifying input is to make it easier for learners to comprehend. When learners can understand what is being said by the interlocutors, it will be easier for them to construct their second language grammars. The following example shows how a teacher of kindergarteners modify their speech based on the addressees.

Example 1: Modified English input instructions in a kindergarten class (Kleifgen, 1985, as cited in Gass & Mackey, 2020)

a. To a group of English NSs: These are babysitters taking care of babies. Draw a line from Q to q. From S to s and then trace.

b. To a single NS of English: Now, Johnny, you have to make a great big pointed hat.

c. To an intermediate-level speaker of English (native speaker of Urdu): No her hat is big. Pointed.

d. To a low intermediate level speaker of English (native speaker of Arabic): See hat? Hat is big. Big and tall.

e. To a beginning level speaker of English (native speaker of Japanese): Big, big, big hat.

From the example, it can be seen that speakers often make modifications in order to make the speech more comprehensible depending on the addressee(s). Simplification, as can be seen from the example above, is not the only way to make adjustments. Modification of speech can also include elaborations. The following example presents a conversation between a native speaker (NS) and a non-native speaker (NNS) in which the NS responded with elaboration when the NNS showed lack of understanding.

Example 2: Elaboration (Gass & Varonis, 1985)

NNS: There has been a lot of talk lately about additives and preservatives in food. In what ways has this changed your eating habits?

NS: I try to stay away from nitrites.

NNS: Pardon me?

NS: Uh, from nitrites in uh like lunch meats and that sort of thing. I don't eat those.

Interactionist research mainly centers on the effects of input on comprehension and L2 development. Some research has pointed out the benefits of interactionally modified input on L2 comprehension (e.g., Pica, Young & Doughty, 1987; Loschky, 1994). This type of input has also been suggested to promote L2 acquisition better than unmodified input (e.g., Mackey, 1999). Although interactionally modified input has been generally recognized as a better alternative, a task-based study on vocabulary learning by Ellis and He (1999) found no difference between pre-modified and interactionally modified input.

1.3.1.2. Negotiation for meaning

According to the interaction hypothesis, negotiation of meaning has a central position in improving learner comprehension and L2 development particularly during a breakdown in communication (Long, 1996). During a conversation between L2 learners and their interlocutors, negotiation of meaning can be identified through its key elements, which consist of clarification requests, confirmation checks, and comprehension checks, all of which signal a communication breakdown (Loewen & Sato, 2018). These elements have been the focus of many research studies which investigate this particular discourse move (e.g., Ellis, Basturkmen & Loewen, 2001a; Loewen, 2004; Gass, Mackey & Ross-Feldman, 2005).

The first element of negotiation of meaning is confirmation checks. It is usually performed when interlocutors need to ensure whether they have correctly understood what has been said. It can be in the form of repetition of the questioned utterance with rising intonation, or a question ‘do you mean X’ (Loewen & Sato, 2018). In the following example, two learners are discussing the objects in the pictures at hand during a spot-the-difference task. Learner 2 checks to confirm whether she correctly understood the information that has been provided by Learner 1, to which Learner 1 responds affirmatively.

Example 3: Confirmation check (indicated by SMALL CAPS) (Gass et al., 2005: 585)

Learner 1: *En mi dibujo hay un pajaro*. 'In my drawing there is a bird.'

Learner 2: *¿SOLAMENTE UN? Tengo, uh, cinco pajaros con un hombre, en sus hombros.* 'ONLY ONE? I have, uh, five birds with a man, on his shoulders.'

Learner 1: Oh, oh, sí, sí. 'Oh, oh, yes, yes.'

The next element of negotiation of meaning is the clarification request. It is defined as an attempt to get extra information from the interlocutor regarding the meaning of what they have said, usually using questions such as "What do you mean?" (Loewen & Sato, 2018). In the following situation, which occurred during an information and opinion task, it can be seen that Learner 2 seeks for more information from his interlocutor using a simple question "What?"

Example 4: Clarification request (indicated by SMALL CAPS) (Gass et al., 2005: 586)

Learner 1: *¿Qué es importante a ella?* 'What is important to her?'

Learner 2*: ¿COMO?* 'WHAT?'

Learner 1: *¿Qué es importante a la amiga? ¿Es solamente el costo?* 'What is important to the friend? Is it just the cost?'

The last main component of negotiation of meaning is comprehension checks, which is usually done to confirm whether an utterance has been correctly understood by the addressee (Loewen & Sato, 2018). Questions such as "Do you understand what I said?" or "Is it clear?" are usually used in this situation. In the following example, Learner 1 asks whether Learner 2 wants her to repeat what she has said.

Example 5: Comprehension check (indicated by SMALL CAPS) (Gass et al., 2005: 586–587)

Learner 1: *La avenida siete va en una direccion hacia el norte desde la calle siete hasta la calle ocho.* ´*¿QUIERES QUE REPITA?* 'Avenue Seven goes in one direction towards the north from Street Seven to Street Eight. DO YOU WANT ME TO REPEAT?'

Learner 2: *Por favor.* 'Please.'

1.3.1.3. Negotiation of form

It is true that negotiation for meaning regularly occurs during communication. However, it has been observed that this type of negotiation does not occur in high frequency in the classroom context (Foster, 1998; Eckerth, 2009). In classrooms, where teachers have a prominent role in interaction, there is an additional type of negotiation that commonly occurs, namely negotiation of form. Negotiation of form generally takes place as a result of a need for linguistic accuracy due to teachers' pedagogical intervention (e.g., Lyster & Ranta, 1997; Ellis et al., 2001a; Lyster et al., 2013). Compared to negotiation of meaning, which occurs due to communication breakdown, negotiation of form has a more didactic function (Lyster, 1998: 190) which oftentimes contains corrective feedback. When a learner produces a linguistically problematic utterance, the teacher usually responds with corrective feedback that is didactic (e.g., didactic recasts). The following example shows a learner using the wrong preposition to which the teacher responds with corrective feedback.

Example 6: Corrective feedback (indicated by SMALL CAPS) (Loewen 2005: 371)

Will:	when I was soldier I used to wear the balaclava
Teacher:	and why did you wear it Will, for protection from the cold or for another reason
Will:	just wind uh protection to wind and cold
Teacher:	PROTECTION FROM
Will:	uh from wind and cold
Teacher:	right, okay not for a disguise

A large number of studies on corrective feedback have been done in the past two and a half decades (e.g., Lyster & Ranta, 1997; Long, Inagaki & Ortega, 1998; Ammar & Spada, 2006; Ellis, Loewen & Erlam, 2006; Mackey, 2006; Yang & Lyster, 2010; Li, Zhu & Ellis, 2016; Nakatsukasa, 2016), which have allowed for many research syntheses (e.g., Long, 2007; Lyster, Saito & Sato, 2013; Nassaji, 2013; Ellis, 2017) and meta-analyses (e.g., Russell & Spada 2006; Li 2010; Lyster & Saito, 2010; Brown, 2016). From these studies, several distinctions of corrective feedback have been documented based on their nature, such as (a) negative and positive feedback (Leeman, 2003), (b) input-providing and output-prompting (Lyster, 2004; Goo & Mackey, 2013; Lyster & Ranta, 2013), and (c) explicit and implicit feedback (Sheen & Ellis, 2011; Lyster et al., 2013). Negative feedback can be identified when interlocutors provide learners with an indication that their utterance is not acceptable according to the standard of the L2. In contrast, positive feedback is when interlocutors show the learners examples of the correct forms directly without telling them that their utterances are not correctly formed (Loewen & Sato, 2018). Several studies have pointed out the positive effects of these two types of feedback on L2 learning (e.g., Schachter, 1991; Leeman, 2003).

Similar to positive feedback in the first distinction, input-providing feedback is done by giving learners the correct linguistic form for the learner. For instance, when learners produce an incorrect utterance, the interlocutors can provide them with the correct form directly after the learners' utterance. An example of this is a recast i.e., a reformulation of the learners' incorrect utterance immediately after they produce it (Loewen & Sato, 2018). On the other hand, output-prompting corrective feedback, instead of providing the correct form, stimulates learners to produce the correct form by themselves. There have been some arguments on which type of feedback is more effective. Some support the use of input-providing feedback (e.g., Long, 2007; Goo & Mackey, 2013) while others support output-prompting feedback (e.g., Lyster 2004; Lyster & Ranta 2013). However, some studies have reported similar effects between the two leading to the suggestion that teachers should use a variety of feedback types on their learners (Loewen & Nabei, 2007; Lyster & Ranta, 2013; Ellis, 2017).

Another issue that has been discussed is whether feedback should be explicit or implicit (Lyster et al., 2013). Some argue that implicit feedback such as a recast is more preferable because it minimizes any interruption (e.g., Long, 1996, 2007; Goo & Mackey, 2013). Long, (2015) himself argues that implicit negative feedback 'does the job' which then allows students and learners to focus on 'tasks and subject-matter learning'. On the

other side of the argument, some researchers (Lyster, 2004; Ellis, Loewen & Erlam, 2006; Loewen & Philp, 2006; Lyster & Ranta, 2013) believe that explicit feedback is more effective because it can be easily recognized by students, which consequently allows them to evaluate their target language repertoire.

Another example of negotiation regarding linguistic accuracy that occurs during communication is called language-related episode (LRE). Swain & Lapkin (1998: 333) state that, during an LRE, interlocutors 'generate [linguistic] alternatives, assess [linguistic] alternatives, and apply the resulting knowledge to solve a linguistic problem'. While engaging in communication, learners sometimes discuss specific linguistic items, even though the communication mainly focusses on meaning. Researchers have acknowledged that an LRE during interaction can serve as a learning opportunity for the interlocutors (e.g., Swain & Lapkin, 1998; Williams, 2001b; Storch, 2002; Loewen, 2005; Kim & McDonough, 2008; Garcıa Mayo & Azkarai, 2016). The following is an example of LRE showing cooperative interactions on a linguistic issue. It can be noticed from the example that corrective feedback is not always necessary in LRE.

Example 7: Language-related episode (Fernandez Dobao, 2016: 40, as cited in Gass & Mackey, 2020)

Larry: entre dos rascacielos, grandes 'between two big skyscrapers'

Ruth: dos 'two'

Jenny: qu´e es? 'what is it?'

Larry: skyscrapers

Jenny: rascacielos? 'skyscrapers?' oh!

Ruth: rascacielos rascacielos 'skyscrapers skyscrapers'

Jenny: look at you

Larry: s´ı 'yes'

Jenny: rascacielos 'skyscrapers'

Ruth: okay

1.3.1.4. Output

Output is the language that is produced by learners during interaction. Swain (1985, 1995, 2005) claims, through her Comprehensible Output Hypothesis, that output

not only represents L2 development but is also a 'causal factor' for L2 development in a number of ways. Firstly, she argues that when learners produce an utterance in L2, they have to think through which forms encode which meanings. This means that they tend to have a greater awareness of the forms of their L2 production (i.e., the noticing function) compared to when they process utterances from an interlocutor. Moreover, Swain argues that through output, learners may test their linguistic hypothesis through feedback that they may receive from the interlocutors (i.e., the hypothesis-testing function). For instance, after learning about a particular L2 structure, a learner decided to try it out during a communication task. During which, they often used it incorrectly. Shehadeh (2001) used the term trigger to refer to the trouble source produced by one of the interlocutors during interaction. Interlocutors may or may not react to it. When they ignore the trigger, it is impossible for the researcher to assume that a breakdown in comprehension or communication has occurred (Shehadeh, 2001). However, the ongoing discourse may indicate whether the listener has not understood or that the speaker ran into difficulty but did not initiate self-correction (Hawkins, 1985; Varonis & Gass, 1985). Alternatively, the listener may react to the trouble source (i.e., negative feedback in the form of recast, clarification request, or explicit correction) or the originator of the trigger may do so (i.e., self-initiated modified output). The outcome can be in various forms. The originator of the trigger may fail to repair, expressing difficulty in repairing or communicating the intended meaning, repeating the trigger without any modification, switching the topic, or successfully reprocessing and reformulating the trouble-source utterance. Swain (1985) argues that SLA is promoted when learners are given more chances to be involved in the negotiation of meaning and this happens when learners can identify the trouble source and successfully modify the output during interaction. This process may cause the learner to revise his or her original hypothesis about the L2 structure. Furthermore, according to Swain, output also has a metalinguistic function which enables learners 'to control and internalize linguistic knowledge' (Swain, 1995: 126). Lastly, since output requires language use by learners, it helps them practice, which can develop fluency and automaticity in L2 (see Lyster & Sato 2013; DeKeyser 2017a).

1.3.1.5. Attention

Attention is the final construct of interaction. It is cognitive in nature, whereas the previously discussed constructs are more discoursal (Loewen & Sato, 2018). Long (1996) argues that interaction 'connects input…; internal learner capacities, particularly selective

attention; and output…in productive ways' (451–452). The importance of attention in L2 learning has been supported by many. Schmidt (1990, 1995, 2001), with his noticing hypothesis, claims that L2 learners need to notice linguistic features in the input that they are exposed to in order to internalize those features. Correspondingly, Robinson (1995, 1996, 2003) believes that attention is indispensable in L2 learning. Attention, according to him, is the 'process that encodes language input, keeps it active in working and short-term memory, and retrieves it from long-term memory' (2003: 631).

As the key constructs of interaction have been identified, researchers are now particularly interested in investigating how these constructs, especially negotiation for meaning, corrective feedback, and output, are affected by the characteristics of the interlocutors, characteristics of the tasks, linguistic targets, and the contexts in which they occur (e.g., Li, 2010; Lyster & Saito, 2010; Plonsky & Gass, 2011; Mackey et al., 2012; Goo & Mackey, 2013; Lyster & Ranta, 2013; Plonsky & Brown, 2015; Ziegler, 2016; Kim, 2017).

1.3.2. Interlocutor characteristics

1.3.2.1. The status of L1

One of the main interests in interactionist research is how interaction is carried out between L2 learners and L1 speakers (or NS) and other L2 speakers (or NNS) (see Long & Porter, 1985). Researchers are particularly interested to find out whether interactions between NS and NNS or NNS and NNS contain constructs that support L2 learning such as input modifications and corrective feedback (e.g., Pica, 2013). Many studies on this topic are carried out mainly in laboratory settings since not many L1 speakers are available in L2 classrooms apart from the teacher (Loewen & Sato, 2018). Moreover, there have not been many studies to investigate L2 learner interactions that occur naturally in L2 contexts (Pérez-Vidal, 2017).

Existing studies comparing interactions between NS-NNS and NNS-NNS mainly focus on four constructs of interaction: input modifications, corrective feedback, modified output, and self-initiated modified output (Loewen & Sato, 2018). In terms of input modification, some studies have found that as input providers, NS are more likely to produce richer vocabulary and more complex sentences when compared to NNS (e.g., Pica et al, 1996). Pica et al (1996) compares NS-NNS and NNS-NNS interaction in two information gap tasks and found that NS tend to provide more lexical and morphosyntactic modifications in one of the tasks. However, a similar study by Garcıa

Mayo and Pica (2000) found that advanced L2 speakers can also provide a richer input than NS. Therefore, a presence of advanced L2 speakers in a classroom (e.g., NNS teacher) as one of the interlocutors can provide comparably similar input to that of NS. Sato (2015) in a more recent study found that even L2 learners can provide a comparable density and complexity in their speech production to that of NS, mainly due to the linguistic simplifications that they tend to produce. However, it is noticeable that the learners sometimes produce input that is grammatically incorrect and solve communication breakdown during interaction using non-target-like solutions (Sato, 2015; Loewen & Sato, 2018).

In terms of feedback, researchers are mainly interested in two aspects i.e., learners' signalling of non-understanding and learners' provision of feedback (Loewen & Sato, 2018). As for the first aspect, the aforementioned study by Pica et al. (1996) shows that, during interaction, learners tend to be more willing to indicate a lack of understanding to another learner than to an NS. They concluded that learner-learner interaction 'did offer data of considerable quality, particularly in the area of feedback' (Pica et al, 1996: 80). Eckerth's (2008) study on learner-learner interaction supports this conclusion, finding that the learners in his study provided their peers with 'feedback rich in acquisitional potential' (Eckerth, 2008: 133) on both targeted and incidental linguistic structures.

Some studies also reveal that L2 learners tend to react more to feedback by revising their problematic structure (i.e., modified output) when they are interacting with their peers compared to NS. This modified output, however, is scarcer during learners' interaction with NS. For example, a study by Sato and Lyster (2007) found that Japanese learners of English modified their problematic utterance more often after they received feedback from their peers than when they received feedback from NSs. Mackey, Oliver and Leeman (2003) supported this claim with their research involving 24 lower-intermediate learners of English from different L1 backgrounds and L1 speakers using information gap tasks. The results suggested that while learner-learner pairs produced more output-promoting feedback, there is a similar quality in terms of modified output in both learner-learner pairs and learner-L1 speaker pairs. Another similar study was conducted by Shehadeh (1999) who compared the interactions between L2 learners and between L2 learners and L1 speakers. The findings of the study suggested that L2 learners tend to 'make an initial utterance more accurate and/or more comprehensible to their

interlocutor(s)' (1999: 644) they receive feedback from their peers than from L1 speakers. This tendency also grows when the duration of interaction is extended.

The last construct, which is less investigated, is self-initiated modified output. Research on this construct has indicated that learners tend to self-correct more when they interact with their peers compared to when they interact with L1 speakers (Loewen & Sato, 2018). Self-initiated modified output, or sometimes simply referred as self-corrections, is thought to be 'overt manifestations of the monitoring process' (Kormos, 2006: 123). It is hypothesized that self-corrections can facilitate L2 processing in the same way as modified output as a result of feedback (de Bot, 1996). Shehadeh (2001), who re-examined the data from his previous study (1999), concluded that self-corrections leading to modified output appear to be noticeably higher in frequency during peer interaction than L2-L1 interaction. McDonough (2004) examined interaction among L2 learners and found that learners tend to produce more initiated modified output than to modify their output as a result of feedback from their peers. The findings from these studies suggested that increased peer interaction leads to improved production of some target language features.

While a large number of previous studies compare L1-L2 interaction with L2-L2 interaction, Bowles, Toth and Adams (2014) contributed a new view by involving heritage language (HL) learners. HL learners are defined as learners who have been exposed to the target language at home (Loewen & Sato, 2018). Bowles, et al. (2014) found in their study that HL-L2 peer group interaction had a better potential to reach target-like outcomes than L2-L2 peer group. They also found that there was more evidence of LRE with the first group. Moreover, they suggested that the discrepancy in proficiency between HL learners and L2 learners actually benefits L2 learners more. Finally, they observed that HL-L2 peer group inclined to stay in the target language during interaction compared to L2-L2 peer group.

To sum up, although it has been suggested that L1 speakers can provide a richer exposure of the target language to L2 learners, it does not necessarily mean that interaction with them is better than with L2 peers. In fact, the aforementioned studies have revealed that L2 speakers can even become better interlocutors that promote L2 acquisition. Long and Porter (1985) suggested that this is something that teachers should consider in their classrooms especially for interactive tasks. In addition, Loewen and Sato (2018) suggested that this is good news for teachers since L1 speakers are clearly not always readily available in most L2 classrooms.

1.3.2.2. Peer interaction

Another topic that has been widely studied in interaction, especially in interactional contexts, is L2 learners' interaction with the teacher and with their peers. In classroom settings, this topic becomes vital since classroom interaction is commonly directed by teachers with peer interaction usually occurring during small group activities or communication tasks. Therefore, it is important to understand the differences between these two groups in an instructional context.

The necessity for peer interaction has been acknowledged for several decades. In 1985, Varonis and Gass (1985) suggested that peer interaction provides as 'a good forum for obtaining input necessary for acquisition' (p. 83). Peer interaction has been thought to be the most common type of interaction in communicatively oriented classroom (Loewen & Sato, 2018). In such classrooms, teachers usually utilize task-based language teaching to promote peer interaction. Consequently, many studies have attempted to examine whether this type of interaction can also be helpful in promoting L2 learning.

Peer interaction has been reported to have positive psycholinguistic impact. Philp, Adams, and Iwashita (2014) maintained that peer interaction provides learners with 'a context for experimenting with the language' (p. 17). This is due to the nature of peer interaction, which is relatively longer in period. Therefore, this type of interaction may extend the opportunities for learners to practice the L2, which consequently allows for more time for input and output. From a psychological point of view, peer interaction makes learners more comfortable in processing the target language through error recognition, which results in more feedback and output modifications (Loewen & Sato, 2018). Consequently, overall language production is increased, which provides more chance for the learners to practice the target language. Philp et al. (2014) also added that peer interaction is less stressful than teacher-led interaction because learners do not feel watched. Learners in Sato's (2013) study explained that, in peer interaction, they feel more comfortable because they did not have to worry about making errors with their peers as they do their teachers.

When studying peer interaction, one should also consider the social context. Tomita and Spada (2013) studied classroom interaction of Japanese learners of English. They found that learners sometimes hesitated to speak English in a conversation task because they feel that using English is seen as showing off. This social stigma is quite prevalent in the context of Japanese learners. Yoshida (2013) support this finding in his study of Japanese learners in Australia. Although the learners in the study knew that they

can improve by participating in interaction, they were still hesitant because they were afraid of making mistakes. Therefore, it has been suggested by Storch (2002) that learners' level of comfort may still depend on the relationship within the group of learners. In addition, Sato (2016) also emphasizes that positive mindset of learners towards peer interaction could help them engage in fruitful interaction.

However, not all believe that peer interaction is always beneficial. Some researchers believe that corrective feedback from peers can be poorer in quality compared to feedback from the teachers. Adams (2007) examined adult ESL learners' interaction and found that peer interaction facilitated learning of overall linguistic structures. However, in a different study, Adams, Nuevo and Egi (2011) found that provision of explicit corrections and the development of past tense were significantly negatively correlated. They even concluded that 'feedback may not play as important a role in learner–learner interaction as it plays in native speaker–learner interactions' (2011: 56). In addition, a recent laboratory-based study of 40 university EFL learners in China by Xu, Fan, and Xu (2019) found that learners were hesitant in providing corrective feedback to their peers. The learners in the study also preferred using recasts instead of prompts and explicit correction in their infrequent corrective feedbacks. Moreover, Xu et al. (2019) found that the learners provided more corrective feedback on morphosyntactic errors than lexical and phonological errors.

However, the disadvantages of peer interaction can still be minimized through some interventions. In a classroom-setting experiment by Sato and Lyster (2012), learners were trained on how to notice errors and to give feedback prior to interaction. This was done to minimize infrequent, inaccurate and unfocused feedback that is common in peer interaction. The results show that this intervention improved grammatical accuracy in learners' production. Sato and Lyster (2012) emphasized that learners need to realize that feedback is beneficial for both the provider and the receiver. A more recent study by Dao (2020) also suggested that interaction strategies need to be instructed to learners prior to interaction tasks. The results of Dao's study showed that the interaction strategy instruction generates more idea units, LREs, talk encouragement and reflection, as well as positive emotions among the participants.

1.3.2.3. The role of L2 proficiency

The proficiency level of interlocutors plays an important part in an interaction (e.g., Yule & Macdonald 1990). The studies in this particular area commonly examine

the effects of learners' L2 proficiency on interaction patterns and on L2 learning. For example, Kim and McDonough (2008) found that learners with different proficiency levels promote constructive interaction. They examined peer interaction of Korean learners and found that the pairs with different proficiency levels resolve communication breakdowns better than same-level pairs especially during lexical LREs. However, in an earlier study, Yule and Macdonald (1990) suggested that there is one condition for a prolonged and successful interaction between learners of different proficiency levels i.e., lower proficiency learners needed to be more dominant in the interaction. In another study by Watanabe and Swain (2007), the same learners were paired with both higher and lower proficiency level partners. They analyzed the quantity and quality of LREs in their interaction and found that the degree of collaboration between interlocutors has more effect on the learning outcome than the interlocutors' level of proficiency. Likewise, Storch and Aldosari (2013) found a similar effect on their study and concluded that un addition to their proficiency difference, the relationship of the interlocutors during the interaction needs to be taken into consideration.

There have also been mixed results in studies of interaction between learners with the same proficiency level. They generally compare low proficiency pairs versus high proficiency pairs. Williams (2001a) examined the frequency and types of LREs between these two groups and found that higher proficiency pairs produced more LREs and performed better on posttests, which means they also received more benefits from the LREs. Williams believed that higher proficiency learners have better monitoring of the target language forms. Correspondingly, Nassaji (2013) compared beginner, intermediate, and advanced pairs and found that advanced pairs gained more from focus-on-form episodes (FFEs, similar to LREs).

In contrast to the previously mentioned studies, Iwashita (2001) found no effect of proficiency on interaction. Her study of adult learners of Japanese found no significant difference between pairs with different proficiency levels in terms of the number of corrective feedback and modified output. In addition, the findings in Oliver's (2002) study are in contrast with Williams' (2001b) finding that negotiation of meaning occurred more in lower proficiency pairs. Oliver even claimed that the lower the proficiency, the more clarification requests and confirmation checks occurred. He added that interaction between lower proficiency pairs allows for 'a greater chance that communication breakdown will occur and, hence, a greater need for the use of negotiation strategies' (2002: 107). Similarly, Sato and Viveros (2016) found that learners from lower

proficiency groups had more language-related collaboration than higher proficiency groups. They claimed that the difference was not instigated by the interlocutors' proficiency levels but rather the degree of collaboration. This claim is similar to those of Watanabe and Swain's (2007) and Storch and Aldosari's (2013) which were mentioned earlier.

1.3.2.4. Individual differences

To date, several studies have been conducted to investigate the role of individual differences that can influence interaction (e.g., DeKeyser, 2012; Mackey et al. 2012; Li 2017). However, five individual differences have become the central focuses of these studies namely (a) anxiety, (b) cognitive abilities, including language aptitude and working memory, (c) willingness to communicate, (d) learner beliefs, and (e) age (Loewen & Sato, 2018).

There have been some reports that anxiety could affect interaction. Anxiety has been defined by Horwitz, Horwitz and Cope (1986) as 'the subjective feeling of tension, apprehension, nervousness, and worry' (1986: 125). They used the term communication *apprehension* as a type of anxiety state which may hamper learners' interaction by deterring their ability to process input and produce output (Sheen, 2008). Although many have studied the effects of anxiety on L2 learning (Dewaele, 2017), only a few have examined the implications of L2 anxiety from an interactionist perspective. One of the earliest examples of such studies includes Sheen's (2008) research, which examined anxiety and corrective feedback. Sheen found that learners with a lower level of anxiety gained more benefits from feedback compared to those with high level of anxiety. In addition, Rassaei (2015) suggested that EFL learners with low anxiety process corrective feedback better than those with high anxiety.

In contrast to the aforementioned findings, there are several studies which found no effects of anxiety on interaction. In one study, Révész (2011) assessed learners' anxiety using questionnaires and examined the effects of the anxiety on learners' L2 production during tasks with different levels of difficulty. Unexpectedly, he later found that there was no significant impact of high anxiety on task performance. Another study by Baralt and Gurzynski-Weiss (2011) investigated the effects of interaction context on L2 anxiety in L2 Spanish learners. They compare the levels of anxiety of the learners when they participated in one-on-one task-based interaction with their instruction using two different modalities. One group of learners performed the task face-to-face (FTF) while the other group used computer-mediated communication (CMC). Learners were

asked to fill in an anxiety questionnaire halfway through the task and also in the end of the task. The results of their study showed no differences in the learners' anxiety levels between modalities nor in the different times of anxiety assessment.

In recent years, there has been an increasing interest on cognitive differences among interactionist researchers. The constructs that have particularly received more attention are language learning aptitude and working memory. Language learning aptitude has been defined as a set of cognitive abilities that are 'predictive of how well, relative to other individuals, an individual can learn a foreign language' (Carroll & Sapon, 2002: 23). A considerable amount of literature on language learning aptitude has been published in the area of SLA (e.g., Carroll, 1981, 1990). However, only few have been conducted in an interactionist perspective (e.g., Mackey et al., 2002; Goo, 2012; Révész, 2012). These few studies have suggested that a higher aptitude can be beneficial for interaction. For example, Li (2013) investigated the effects of cognitive differences on the effectiveness of corrective feedback and Chinese classifiers. Li found that language analytic ability was a significant predictor of test gain scores which he concluded that 'in the absence of metalinguistic information, learners with higher analytic ability achieved more' (2013: 647). In another study, Trofimovich, Ammar and Gatbonton (2007) investigated learners' working memory, phonological memory, analytical ability, and attention control and correlated them to the learners' ability to notice and benefit from recasts. In general, the learners in their study could notice and benefit from recasts with substantial individual variability. Attention control was the strongest predictor for the gain scores with others being phonological memory and analytical ability.

The next element within the construct of cognitive ability is working memory. It has been generally assumed that there is a positive correlation between learners' working memory capacity and the benefits that they will gain from L2 interaction. Révész (2012) studied EFL learners in Hungary and suggested that high phonological short-term memory helped learners to benefit from recasts. In the same study that was mentioned previously, Li (2013) investigated whether working memory, assessed through a listening span test, promotes the effectiveness of explicit and implicit feedback. The results of his study showed that learners with working memory performed significantly better on an elicited imitation test and grammatically judgment test. Finally, Kim, Payant, and Pearson (2015) investigated whether task complexity and working memory have any impact on learners' noticeability of recasts and their question formation ability. The study found that

only working memory has a significant impact on both noticing ability and the L2 development.

Despite the growing number of studies on the relationship between psychological individual differences and interaction, there are still some gaps that need to be filled in some areas of the field. One instance of these gaps is the relationship between motivation and interaction, which is largely still under-investigated (Dornyei, 2002). In this case, one type of motivation, which is called task motivation, has also been overlooked by researchers in the area (see Dornyei & Kormos, 2000). This type of motivation describes 'why students behave as they do in a specific learning situation where they are carrying out a specific task' (Csizer, 2017: 424–425). Another construct that could impact learners' interaction behavior but has been overlooked is willingness to communicate (WTC) (see MacIntyre, Burns & Jessome, 2011). MacIntyre et al. (1998) described it as 'the probability that a learner will use the language in authentic interaction' (1998: 558). Cao and Philp's (2006) study is one of few. They investigated whether learners' self-reported WTC has any impact on their classroom behavior. However, the results of their study indicated little correlation. In a subsequent study, Cao (2014) suggested that WTC should be regarded more as a construct that is dependent on 'dynamic situational variables' (p. 789) instead of as a part of individual differences.

In terms of learners' beliefs, the hypothesis is that if learners respect interaction as an activity that will benefit them, then they may benefit more from interaction. On the other hand, if learners believe that grammatical and vocabulary drills should be applied in their classroom instead of interactional tasks, then they may not benefit from interaction. Schulz (2001) investigated teachers and learners' beliefs on grammatical instruction and error correction and found that learners prefer error correction, while the teachers valued grammar instruction more. According to Schulz, this situation may negatively impact language learning unless the learners' beliefs are facilitated. In a large-scale study in the US involving 700 learners studying different languages, Loewen et al. (2009) found that ESL learners valued interaction more compared to the learners of foreign languages. Such positive beliefs of the ESL learners may make them more receptive towards interaction activities from which, as a result, they could gain more benefits. Loewen and Sato (2018) emphasized that learner beliefs may be changed by teachers' instruction. Sato (2013) demonstrated this is an intervention which was aimed at raising learners' awareness on the benefits of peer interaction. Through this intervention, the learner beliefs became more positive towards interaction, represented in

the increased amount of corrective feedback produced by the learners in the later interaction. Together, these studies indicate the positive correlation between learners' beliefs on interaction and the benefits that they will gain from it.

The next individual difference that will be discussed here is learners' age. Age has been considered as one of the most influential individual differences in instructed second language acquisition (ISLA), which can be seen from the fact that teachers may adjust their teaching strategies based on their learners' age groups. While a large and growing body of literature has investigated the effects of age of acquisition on ultimate attainment (see DeKeyser, 2017b), only a few studies on the relationship between age and interaction are available due to some methodological complications (see Oliver, Nguyen & Sato 2017). Nonetheless, the available studies (e.g., Oliver, 1998; Oliver, 2000; Oliver & Mackey, 2003; Mackey, Kanganas, & Oliver, 2007; Oliver, Philp & Duchesne, 2017) have indicated that interaction benefits adult and younger learners differently. For example, Oliver (1998) found that child learners produced fewer clarification requests and confirmation checks. In a more recent study, Oliver, Philp and Duchesne (2017) compared two groups of young learners (5-8 years and 9-12 years) and found that, in some topics of the task, the older group of learners tend to produce less negotiation of meaning because they 'simply wanted to get the task done' (2017: 8).

Overall, the studies that have been reviewed make it clear that individual differences may have some impact on learner interaction. However, further investigation is still needed in order to shed more light on this particular topic. There are several individual differences that can still be explored particularly in relation to interaction including personality and learning styles (Dewaele 2017) and individual creativity (McDonough, Crawford & Mackey 2015). As a dynamic construct, learner psychology always has more to offer for researchers interested in this area.

1.3.3. The role of context in interaction

There also have been several studies investigating the characteristics of interaction which occurred in different contexts. Lyster and Ranta (1997) conducted a study in a French immersion high school describing the frequency and types of corrective feedback that the teachers and the learners produced during communicative tasks. They found that teachers produced at least six types of feedback with recasts as the type of feedback that the teachers provided most. Lyster and Ranta also found that learners responded to the feedback differently. Lyster (1998) analysed the same data set further and found that

learners' uptake occurred more when they respond to output-prompting feedback. The studies also suggested that recasts were less effective when the aim is to promote learners' L2 production.

Some studies have investigated the types of corrective feedback and output responses (i.e., uptake) occurring during interaction in New Zealand ESL classes (Ellis et al. 2001a, 2001b; Loewen 2004). The results of the studies suggested that recasts were the most frequent type that occurred in the classrooms. Moreover, these studies also indicated that approximately 75% of the feedback resulted in successful uptake by the learners. Sheen (2004) compared the data from the previous studies with her own data from communication classes in South Korea. Her study indicated that the uptake rates in the studies in Canada (French immersion: 54%; Canadian ESL: 46.6%) were significantly lower than in New Zealand (80.4%) and South Korea (82.3%). She also found that in the South Korean (82.5%) and New Zealand (72.9%) contexts, learners responded significantly more to recasts than in the Canadian context (French immersion: 30.7%; Canadian ESL: 39.8%). Taken together, these studies support the notion that the type of classroom and sociolinguistic context of the target language may have some effect on the response of learners to corrective feedback.

A more recent study in classroom contexts was conducted by Bowles, et al. (2014). They involved L2 learners of Spanish and heritage learners of the language and examined the types of interaction they produced during some communicative tasks. They found no significant difference in the types of interaction whether the interlocutors were from L2 learners or HL learners. In another study, Basterrechea and Garcıa Mayo (2013) investigated the effects of instructional context on LREs in two different settings in Spain; a content language integrated learning (CLIL) setting and EFL setting. They found that LREs occurred significantly more in the CLIL setting than in the EFL setting. Overall, these studies were important examples of purely descriptive studies exploring the occurrence of interaction in classroom settings.

1.3.4. Methods in interactional studies

1.3.4.1. Laboratory and classroom study

It is obvious that naturalistic L2 interaction can also happen outside the classroom context (e.g., Schegloff 2000). However, the main concern of the available research in the area of ISLA has been interaction that takes place in pedagogical context and/or interaction that is, in some ways, manipulated by the researchers (i.e., laboratory context)

(e.g., Mackey 2012; Loewen 2015). It is argued that interactionist research conducted in classroom settings provides high ecological validity since it represents the interaction involving teachers and students that usually happens daily in the real world (Loewen & Sato, 2018). Since it is sometimes difficult for researchers to manage the complexities of the variables in the classroom context (Shadish, Cook & Campbell 2002), laboratory studies have been conducted to serve as an alternative in examining the nature of L2 interaction. In a meta-analysis, Mackey and Goo (2007) found that from 28 studies that they analyzed, 64% of them were conducted in laboratory settings, while the rest were conducted in classroom settings.

However, some have questioned the generalizability of laboratory research (e.g., Foster, 1998; Eckerth, 2009). The high extent of intervention in laboratory research may result in higher awareness of linguistic forms in learners' output. As a result, learners may perform more noticing of linguistic forms, which may also affect their pre-tests or post-tests scores. This may give a false representation of L2 development (Loewen & Sato, 2018). A meta-analysis of 28 studies by Mackey and Goo (2007) suggested that this might be the case. The results of their analysis showed that the effects of interaction on posttests in laboratory research were greater. In contrast, another meta-analysis of 15 studies by Russell and Spada (2006) found no differences between classroom and laboratory setting. Mackey et al. (2013) later suggested that more authentic classroom research is needed particularly in examining the effects of interaction on L2 development. And if we take a dynamic usage based perspective (see Section 1.4), then also long-term effect need to be taken into consideration.

1.3.4.2. Descriptive and quasi-experimental study

It can be noticed from the previously mentioned studies that the early studies of interaction were mainly observational (e.g., Gass & Varonis 1986). These observational studies mainly described interaction as it is happening, while other studies correlated the interactional features to the ability of noticing and the use of L2. However, since more studies have revealed the characteristics of interaction as well as the variables that affect them, more quasi-experimental studies were conducted to manipulate one of those variables to better understand their effects on L2 acquisition (Loewen & Sato, 2018). There are still some recent descriptive studies attempting to explore several features that affect interaction, which were relatively unknown. Bowles, et al. (2014), for example, is one of the first who consider the role of heritage language learners in interaction. In

addition, as communication technology continues to develop, some descriptive studies also take this into consideration by investigating the effects of technology on interaction (e.g., Loewen & Wolff, 2016).

With regard to the studies that examined noticing, several methods have been used and they can be categorized into two groups, namely concurrent and retrospective measures (Loewen & Sato, 2018). When using concurrent measures, researchers assess noticing in real time when an interaction is taking place. For example, some studies of uptake (e.g., Ellis et al. 2001a; Loewen 2005) examined learners' responses to corrective feedback during interaction. They argued that learners' responses may show some evidence of learners noticing the form (Lightbown, 1998). It may also be possible, however, that during interaction, learners may notice the L2 form without any signal or response and vice versa; learners may echo the corrective feedback provided to them without noticing the L2 form. Hence, these studies have admitted that the features of interaction are not always equal to cognitive processes leading to L2 learning. Consequently, researchers have employed other methods of concurrent measures to examine noticing. One example of such measure is by requiring learners to respond to a stimulus at some important stages of the interaction. In some studies (e.g., Philp 2003; Bigelow et al. 2006), learners were asked to repeat the recasts or to report what they were thinking right after receiving the recasts. These measures were taken to prove whether the learners in fact notice the feedback that they had received. Recently, technology has been used to assist researchers in carrying out concurrent measures of noticing. For example, eye-tracking technology has been employed to measure learners' noticing when they are participating in written synchronous computer-mediated communication (SCMC) (Sauro & Smith 2010) and in online FTF interaction (McDonough et al. 2015). This technology is used under the assumption that eye movement could indicate cognitive attention (see McDonough, 2017).

The use of concurrent measures has been criticized because of the interruptions which affect the flow of conversation. Consequently, retrospective measures become more popular in examining learners' noticing during interaction. Gass and Mackey (2000), for example, used stimulated recall, a popular method in retrospective measures. In this method, learners watch or listen to their recorded interaction and are asked about what they were thinking during critical moments in the interaction. Some studies using stimulated recall have revealed that learners noticed L2 lexis and phonology better than morphosyntax (Mackey, et al, 2000). Besides stimulated recall, other examples of

retrospective measures are learner journals and a list of focused questions to get learner responses about the interaction. Mackey (2006) employed all three retrospective measures that have been mentioned to investigate learners' noticing and found some inconsistencies in the learners' report, which led her to suggest that noticing should be viewed as a continuum.

To summarize, many studies have examined the effects of interaction and feedback on L2 development. In the next chapter we will detail the aspects that will be considered in the current dissertation, in which we will look closely at peer interaction among pesantren students.

1.4. Second language development from a dynamic usage-based perspective

1.4.1. Dynamic usage based perspective

As Long (1998) points out there has been an absence of either a widely accepted theory of language learning or a solid empirical base for classroom practice. Much L2 developmental research has taken place without an underlying theoretical linguistic theory of what language is and how it changes, but in SLA research the focus has been primarily on the acquisition of morpho-syntax, probably because that is where the communicative or immersion programs seemed to fall short (Hammerly, 1991) and both researchers and teachers have sought ways to avoid fossilization. However, recently usage-based theories have found their way in SLA, according to Tyler et al. (2018). In their introduction, the editors maintain that there is not one definite usage-based model of language and language learning, but rather a family of approaches that include cognitive linguistics, emergentism, constructionism and complex dynamic systems theory. In this dissertation, we will focus on a dynamic usage based linguistic approach as proposed by Verspoor (2017).

The concept of dynamic usage-based approach is essentially a combination of two existing approaches namely complex dynamic system theory (CDST) and usage-based linguistics (UBL) (see Langacker, 2009; Verspoor & Behrens, 2011; Verspoor, Schmid & Xu, 2012; Roehr-Brackin, 2015). Both theories perceive developmental changes as emerging from language learning as a result of language use. While the two approaches seem similar, they actually have different origins. CDST is originally not a specific theory of language learning but a theory of change (Larsen-Freeman & Cameron, 2008) used in different disciplines such as mathematics, chemistry, physics, etc. On the other hand, the usage-based perspective has a linguistic origin. It comprises various linguistic

approaches, which suggest that linguistic knowledge can be described as a repertoire of constructions with different extents and levels of abstractness (Barlow & Kemmer, 2000; Cadierno & Eskildsen, 2015). Thus, there is some difference in this matter.

L2 development can be regarded as a dynamic process of change (Larsen-Freeman, 1997; De Bot, 2008; and Verspoor, De Bot and Lowie, 2011). CDST identifies a complex system such as language as a group of entities, sub-systems or variables that are interconnected, continuously interact, self-organize and coordinate as a whole. The dynamics of such iterative processes cause changes to be non-linear with a significant extent of variability (within systems) and variation (among systems). As De Bot and Larsen-Freeman (2011) put it, sometimes a system "changes continuously, sometimes discontinuously, even chaotically". Systems, however, have a tendency to move towards preferred states, called attractors. Therefore, language development cannot be represented by a straight linear continuum. The development is obviously complex and elusive, involving numerous dimensions that develop at mixed and non-linear rates.

Thus, in its approach, learner language constructions are perceived as a network in which all the constructions are interconnected. Through the process of learning, the network evolves sometimes gradually and sometimes suddenly. The dynamic usage-based (DUB) approach addresses these changes in learner language which are a consequence of, 1) the frequency of use of L2 in social interaction, and 2) the interaction of constructions in the network in the learner's mind (Verspoor & Behrens, 2011; Roehr-Brackin, 2015). Moreover, in the DUB construct, initial conditions of the learners play a key role in learners' development and, therefore, learners are expected to have different individual trajectories. Learners' personal and linguistic background such as L1, age, scholastic aptitude, motivation, etc. are assumed to serve as variables which interact in complex manners and determine the acquisition of the L2. Furthermore, it is also believed that sub-systems of any organism are in some way interconnected and affect each other continuously in the development process (de Bot, Lowie, & Verspoor, 2007; van Geert, 1991).

Therefore, a DUB perspective has a strong stance against the argument that there are innate structures specific to language. Verspoor et al (2012) argue that language is "intrinsically linked to general cognitive process (interconnected variables in CDST terms) and intrinsically symbolic through form-use-meaning mappings (FUMMs) (coordination in CDST terms), constituted by a structured inventory of linguistic constructions, i.e., conventionalized form-meaning pairings used for communicative

purposes (emergence and attractors in CDST terms) (2012: 67)". They added that FUMMs exist at various levels of complexity and abstraction, consisting of concrete and more abstract classes of items, and complex combinations between concrete and abstract linguistic expressions and, therefore, there is a flexible boundary between lexis and grammar (see Langacker, 2000; 2008; Tomasello, 2003). They stated that language development begins with more concrete items (item-based) and gradually moves towards more abstract linguistic schema in an implicit and inductive process (self-organization in CDST terms). But the most important tenet of all usage-based approaches is that language is learned from the input and output that the language learner experiences. There are no innate systems, so the language learner can only discover and acquire the language through exposure and experience. Frequency of exposure is supposedly the main driver of development. Whatever is heard or used the most will become automated and entrenched patterns in the learner's language.

1.4.2. Language development studies in a DUB perspective

Language development studies using a DUB perspective usually focus on finding out how different variables of the language develop and the way they interact over time (e.g., Verspoor, Schmid & Xu, 2012). There are several studies that have been conducted within this framework. For example, Larsen-Freeman (2006) conducted a repeated task study involving five Chinese learners of English with intermediate proficiency. She assessed their oral and written production, particularly measuring the development of the construct of complexity, accuracy and fluency (CAF) in the learners' L2 production by means of both holistic and specific measures. She found that the use of L2 by the learners influence their language resources. She also discovered that the learners show diverse patterns in the CAF development graphs. In a longitudinal study, Spoelman and Verspoor (2010) also tracked the CAF construct of in the written production of a Dutch learner of Finnish. In over three years, they studied their development from beginner to high intermediate level. Their study reveals that the interaction among the measured variables show "classic jumps, transitions, and non-linear" development. Besides beginner and intermediate learners, these variability patterns and interactions among variables have also been observed in advanced L2 learners (e.g., Verspoor, Lowie & Van Dijk, 2008). In another study, Caspy (2010) modelled a developmental path of four learners and found that three of them developed lexical and syntactic complexity before lexical and syntactic accuracy. This indicates that the learners' lexicon usually develops before the syntax and

complexity before accuracy. This is a very reasonable developmental order. Language learners first need words to make longer and more complex sentences and accuracy will come only after continuous input and practice.

Another study in this theoretical framework by Verspoor et al. (2012) explored which variables are most likely to change at different proficiency levels. The study used written texts to track L2 learners' development from beginner to high-intermediate level of proficiency. Over forty complexity and accuracy variables were measured and the results indicated five general measures which showed almost linear progression and regression across the levels of proficiency, i.e., less simple sentences (sentence complexity), less present tenses (verb phrase complexity), less errors (accuracy), and increased type-token ratio (lexical diversity) and increased instances of chunks (authentic expressions). Other specific measures (less-frequently occurring structures), such as perfect or progressive verb phrases or specific types of dependent clauses, almost all showed variation, non-linear development, and fluctuating relationships. Additionally, they found that learners in the earlier stages experienced significant development in lexical measures, then in syntactic before finally advancing in lexical measures again, particularly in the use of chunks. These findings were supported by other longitudinal studies. For instance, Bulté (2013) who examined the development of L2 complexity of a sub-group of the Verspoor et al. (2012) participants found that the increase in L2 complexity is fairly linear at group level, but he found a great degree of variability at the individual level.

In another study, Hou, Verspoor and Loerts (2016) studied the potential gains in English as L2 development in one group of senior high school students and two groups of university students in China. In a pre-post design, the learners' writing samples were scored holistically and analytically. The analysis of the holistic scores showed that only the senior high school group improved. However, the analytic scores reveal improvements in all groups in different features of the written language. They suggested that different variables may develop when the learners are at different levels of proficiency. From their findings, they recommended that one complexity measure may not be applicable to all proficiency levels.

With regard to individual differences, several studies (usually not from a DUB perspective) have been conducted to examine the role of individual differences in L2 development. Differences including gender, motivation, aptitude, initial proficiency and several other learner characteristics have been reported to affect L2 development. Some

studies (e.g., Oxford, 1993; Young & Oxford 1997), for example, have indicated that gender can play a significant role on how language learners develop. Such studies have generally suggested that females are better language learners because they tend to have more positive attitude towards L2, show better integrative motivation, and utilize a wider range of learning strategies (Oxford, Nyikos & Ehrman, 1988). Motivation has also been widely considered as one of the most prominent factors affecting at L2 acquisition. One of the leading researchers in this area, Gardner (1985), found that higher motivation could result in more desire and effort to achieve learning goals. It also leads to better attitude in the learning process. Saville-Troike and Barto (2016) even claims that motivation is the second most significant predictor after aptitude in the success of second language learning. However, this claim may not always be correct. For instance, in a study conducted by Verspoor, de Bot, and Xu (2015), motivation had a significant contribution in L2 development in one group of participants but not the other.

Another factor that has been reported to affect L2 learning is aptitude. Researchers have proposed two different type of aptitude namely general aptitude and linguistic aptitude. Earlier studies have long confirmed that general aptitude as well as linguistic aptitude play an important role in the success of L2 learning (e.g. Skehan, 1989). Verspoor et al. (2015) confirmed in their study the importance of scholastic aptitude. However, they also argue that from a DUB perspective, L2 development is affected by many different contributors which interact among each other. Therefore, one cannot assume that the success of L2 learning of all learners is due to a single contributor but there are multiple factors playing their roles. In their study, they found that general scholastic aptitude is a strong predictor in one group but not the other.

Finally, a more recent study (Huang, Steinkrauss, & Verspoor, 2021) shows that besides the traditional individual differences, variability has strong correlation with L2 proficiency gain. Variability is an important feature of a dynamic system. Several studies have reported that the degree and pattern of variability can provide an insight into the development of L2 learners (de Bot, Lowie, & Verspoor, 2007; Lowie & Verspoor, 2019; Spoelman & Verspoor, 2010; Rousse-Malpat et al., 2019; Verspoor & de Bot, 2021). Huang et al. (2021) did multiple linear regression analyses and found that variability was a strong predictor of gains and final L2 writing proficiency when the initial proficiency of the participants was controlled. With that being said, this study also considers the degree of variability in the analysis.

To summarize, from a DUB perspective, the most important driver of L2 development is exposure and use, preferably in the form of whole form-use-meaning mappings, and what is used frequently will become entrenched and automated over time. Within the developmental process many factors such as L1, age, motivation, aptitude, and so forth may play a role in development and because there are so many interacting variables development will show variability (ups and downs within a learner) and variation (differences among learners). In the current study we will look at several of these variables and trace learners' L2 development in their written texts during one academic year.

1.5. Pidginization and second language acquisition

A DUB theoretical perspective may inform us about language development that changes over time or stagnates, either with target norms or non-target forms, the latter being called fossilization in SLA studies. Larsen-Freeman (2006) argues that there is no such thing as fossilization as we cannot be sure that the system will not change anymore further down the road as there is not really an end-state. Still, from a DUB point of view, any system may move to an attractor state where it is likely to remain for a longer time, which can be related to studies of pidgin and creole languages. In the current study the term "fossilization" is defined as non-target forms that do not seem to change in the course of one academic year.

The emergence and development of pidgin and creole languages usually involve extreme case of languages in contact, which cause accelerated linguistic change (Lefebvre, 2004). In the beginning, second language acquisition plays a crucial part in shaping the languages followed by first language acquisition in its development when the speakers' children are exposed to the newly created languages.

In 1979, Derek Bickerton and Talmy Givón proposed an experiment in which people speaking mutually unintelligible languages are taught approximately 200 words of English and then placed on an uninhabited island for a year where they would communicate using only the English lexicon while performing agricultural activities (as cited in Master, Schumann, and Sokolik, 1989). Knowing that pidgin languages typically develop under such extreme social condition, the purpose of this experiment was to understand how pidgin languages are formed in their early stages. Eventually, the proposal was rejected due to potential dangers to the subjects. However, their proposal

was vital since it offers an insight into conditions under which a pidgin could be experimentally created and thus opens up possibilities for the study of early pidginization. Bickerton and Givón's proposal provided a basis for Mester et al. (1989) to conduct an experiment to imitate the "Island Experiment" without risking the subjects. In their experiment, they translated the Island Experiment word list into Farsi and German. There were two groups of graduate students involved. In the first group, four participants memorized the words in Farsi while in the second group eight participants memorized them in German. Having memorized the words, each group spent 40 hours using these words to communicate in information gap activities. The results of their experiment, to some extent, reflect the characteristics of genuine pidginization in its early stages i.e., extensive lexical creation, use of "natural" syntax, and lack of grammaticalization.

Mester et al. (1989) conducted the experiments to mimic real life situation in which pidginization process take place. Their experiment and other likewise experiments in applied linguistics are done because the situations that they want are particularly difficult to capture. For instance, pidgin languages usually develop under extreme social condition such as colonial economic systems. Therefore, to find such a condition in real life in order to observe the development of a pidgin language would be unfeasible. However, when such a condition exists and is readily accessible to researchers, it is definitely worth conducting a study there.

Pidgin has been largely described as a contact language which develops when groups of people who speak different languages attempt to communicate with one another (Richards & Schmidt, 2010). Crystal (2010) uses the terms "makeshift", "marginal" language, or "mixed languages" to define pidgin. There is a debate on the etymological origin of the term pidgin, but many have settled on the notion that it is a Chinese corruption of the English word "business" (Baker & Mühlhäusler, 1990 as cited in Bakker, 1994). Pidgins commonly have been observed to have limited lexis, morphology, syntax and a narrow range of use (e.g., trade), which can expand and develop when they are used over an extensive period and when their purposes expand. However, they are not 'bad' versions of the source languages but rather highly regularized varieties (see Todd, 1974; Mühlhäusler, 1986; & Romaine, 1988). Pidgins commonly have no native speakers but some expanded pidgins, e.g., Tok Pisin in Papua New Guinea and Singlish in Singapore, are spoken as first or primary language by some people in the area and are considered as a creole.

The characteristics of pidgins, i.e., simplifications and transfers of the source languages, are also typical of the language of a learner. Nemser (1974) described these features of learner talk as approximative system while Selinker (1972) discussed them in his famous paper on interlanguage (IL). This form may be characterized by a limited system of auxiliary verbs, simplified question and negative forms, and reduced rules for tense, number, and other grammatical categories. For example, some learners may omit inflections in numeral phrases such as "two cat", which is similarly found in the pidginization processes, baby talk (BT) (Ferguson, 1977) and foreigner talk (FT) (Ferguson, 1975). Another example of simplification is auxiliary deletion as in "he in school".

The simplifications and L1 transfers produced by language learners are typically temporary in the language learning process. L2 learners will usually move towards more target like forms in their L2 production as they progress. For example, in a cross-sectional usage-based study on Dutch learners of English, Vries and Verspoor (2010) found that learners' L1 transfer errors go rapidly down between level 1 and 2. The decrease in L1 transfer errors is a common case in the language learning process. However, in some cases, some errors remain and fossilize. Richards (1974) argued that a fossilized form of the target language could be considered pidginized when learners do not advance beyond this stage.

Structural simplifications in both IL and pidgin languages result from the situation of language contact. Richards (1974) closely compared pidgin languages and second language acquisition (SLA). He argued that both codes can be described "as an IL arising as a medium of communication between speakers of different languages, characterized by grammatical structure and lexical content originating in differing sources, by unintelligibility to speakers of the source languages and by stability" (Richards, 1974: 77). Some research (e.g., Cancino, Rosansky & Schumann, 1974; Schumann, 1978; and Andersen, 1981) indicate further the similarities between the structures of pidgin languages and the IL (see Selinker, 1972; 1992) of L2 learners.

Probably the most fundamental framework on the resemblance of IL and pidgin languages was proposed by Schumann (1978) in his acculturation hypothesis. Schumann and his two colleagues (Cancino et al., 1974) initially studied six Spanish learners of English of different ages in the USA for over a period of ten months. Schumann eventually focused his work on one of the learners, a 33-year-old called Alberto, who showed markedly less improvement than the others.

Alberto showed several characteristics of pidgin languages in his speech such as the lack of inflectional morphology. Although he exhibited fair consistency when using certain morphemes such as plural –s (85%) and irregular past –ed (65%), he failed to show the same consistency in the use of regular past –ed (7%) and inversion (5%). In terms of auxiliary verbs, when the other five learners in the study went through the acquisition of auxiliary *is, am, can, do, does, was, did,* and *are;* Alberto only went as far as *is, am, can,* and *are.* He proceeded to use only those four auxiliaries until the end of the observation period with only auxiliary *is* that was acquired satisfactorily. According to Schumann (as cited in Johnson & Johnson, 1999), another similarity between pidgins and Alberto's speech are:

- both use a single negative marker *no* and have a rule that negation can be expressed through a formula of 'no + verb' as in *"I no see";*
- both lack inversion of subject and verb, e.g., *"Where the paper is?"*;
- both lack auxiliaries, e.g., *"She crying";*
- both lack possessive inflection –s, e.g., *"The king food"*;
- both lack inflectional morphology, e.g., *"Yesterday, I talk with one friend"*; and subject pronouns, e.g., *"No have holidays"*.

This comparative study of pidgin languages and the Alberto's language led Schumann (1978) to the conclusion that Alberto's speech was in fact a pidginized version of English. The prominent case of Alberto eventually gave birth to the acculturation hypothesis or also known as the pidginization hypothesis. Schumann believes that a pidginized form of a language may develop for two main reasons; (a) when learners separate themselves socially and psychologically from speakers of the target language, and (b) when the target language is used by learners for a very limited range of functions (Richards & Schmidt, 2010). Andersen (1981) compared Alberto's English IL and Bickerton's (1977) research on Hawaiian Pidgin English to explore further the similarities between both types of linguistic codes. Supplementing Schumann's idea, he concluded that both codes have the following characteristics:

(1) Reliance on word order rather than inflections for expressing grammatical relations.
(2) Native-language transfer in word order as well as use of English word order.
(3) Sporadic merging of pre-verbal markers which come from lexical verbs promoted to auxiliary status.

(4) A basic pidgin negation.

(5) Lack of inversion in questions.

(6) Preponderance of uninflected verb forms.

(Andersen, 1981, cited in McLaughlin, 1987)

Pidgin languages are used primarily for communicating ideas. They are restricted languages that serve only the communicative function; speakers of pidgins normally do not identify themselves with the group who speak the pidgin. They tend to reside in their own group apart from purposes of contact. Alberto had normal intelligence just like the other Spanish learners in the study. However, unlike the other learners, he did not use English for social purposes. In the case of Alberto, the pidginized form of the target language occurs due to social isolation, which is natural in second or foreign language learning. In Schumann's research (as cited in Johnson & Johnson, 1999), the idea of acculturation means social and psychological integration with the target group, which accounts for success and failure of L2 learning. Social factors, on the other hand, are covered by the idea of social distance: when one group dominates the other or when one group isolates themselves from the other etc. If social distance is high then there is little chance for success in language learning. The other factors that are pivotal in the success of language learning are psychological factors. They refer to psychological distance: when the learners experience language shock, or culture shock, or are poorly motivated then the chance for success is low.

Although the idea was intriguing, little research was done to support the theory. Many argued that an L2 theory cannot be based on a single unsuccessful L2 learners (Johnson & Johnson, 1999). Although the theory mainly applies to immigrant situations and not in foreign language situations, Johnson and Johnson (1999) believed that there was a second concept of 'enculturation' that was raised when describing people who learn an L2 in order to function in their own society, e.g., in England or in Russia in the past, a 'gentleman' had to know French, a foreign language which indicates status within a society and has nothing to do with its functions outside the society.

Another case that supports the possibility of a pidgin-like language being produced by students is in the immersion programmes in Canada and the United States (Swain, 1997; Hammerly, 1991). Being critical of this type of communicative approach, Hammerly (1991) especially scrutinized these immersion programmes and concluded that although the students were successful in attaining a high level of communicative proficiency (fluency), they failed in terms of linguistic accuracy. He cites studies which

show that “an error-laden classroom pidgin becomes established as early as Grade 2 or 3 because students are under pressure to communicate and are encouraged to do so regardless of grammar” (1991: 5).

To summarize, there have been several authors who have suggested that the interlanguage in an L2 learner shows similarities to pidginized languages, but the non-target forms may disappear as the learner becomes more proficient. However, the non-target forms may also remain and not change anymore, in which one speaks of “fossilization” in the L2 literature and pidginization if the non-standard language becomes the standard means of communication in a group of speakers who are isolated from the target language speaking community. In the framework of the present study, the notion of pidginization is thus defined as the stabilization of non-target forms in a group of learners who use the language for extensive use in real-life communication only within the group itself.

1.6. Summary and research questions

It is clear that the interaction approach in ISLA has been studied extensively and that proponents believe that learners may acquire the L2 through interaction, but there are differences in who the interlocutors are and what types of feedback may aid language learning the most. The earlier studies played an important role in revealing the characteristics of interaction and consequently enabled interactionist researchers to explore specific variables related to interaction. There are at least three main focuses of interactionist studies, i.e., (a) discourse moves e.g., modified input (Swain, 1985, 1995, 2005), (b) cognitive constructs e.g., noticing (Schmidt, 1990, 1995, 2001), and (c) L2 development and acquisition (Mackey, 1999; Spada & Lightbown, 2009; Mackey, 2012). On the other hand, the most commonly investigated variables have been categorized into four domains: those concerning (a) the interlocutors (e.g., L2 proficiency, L1 status, gender, etc.), (b) the task characteristics (e.g., complexity, type of task), linguistic targets, and (d) the interactional context (Loewen & Sato, 2018).

Some interactionist researchers have suggested that the focus of interactional research has been shifting from investigating whether interaction is beneficial for L2 development to how and under what condition it could be beneficial (Mackey et al., 2012; Pica, 2013; Long 2015; Mackey & Gass 2015; Long 2017). They also have pointed out the need for further research. Mackey et al. (2012) suggested that more replication studies need to be done, particularly to cover the methodological shortcomings which have been

mentioned earlier (Plonsky & Gass 2011). With the complexity of SLA and the dynamics of ISLA, Loewen and Sato (2018) believed that there is always opportunity for replication studies. According to them, there are several variables of interaction which can be explored further, including the benefits of interaction on pragmatics; the role of individual differences; social and sociocognitive issues in interaction; interaction in young learners and 'non-traditional' learners; learners' motivation and engagement; and the roles of gestures in interaction. Besides Loewen and Sato (2018) have pointed out an urgent need for longitudinal studies and delayed testing to understand the long-term effects of interaction. Moreover, there have not been many studies investigating L2 learner interactions that occur naturally in L2 contexts (Pérez-Vidal, 2017). Most interaction research occurred in a classroom or in laboratory settings with the results of the latter generally considered as a reflection of their consequences for the L2 classroom (Loewen & Sato, 2018). However, much less is known about the implication of interaction in naturalistic settings on the development of L2 learners. Thus, more studies are still needed in order to "further our understanding of the effect of interaction on L2 development" and "help extend the parameters of the interaction approach" (Loewen & Sato, 2018: 317).

With regard to peer interaction, the benefits of peer-interaction in L2 learning have been endorsed by many previous studies, which suggest that it provides a good medium for learners to obtain input (e.g., Varonis & Gass, 1985; Eckerth, 2008) and has positive psycholinguistic impact (Sato, 2013; Philp et al., 2014). However, not all studies of peer interaction have supported these notions, noting the lack of quality especially in terms of corrective feedback in peer interaction (Adams, 2007; Adams, Nuevo & Egi, 2011; Xu, Fan & Xu, 2019). As elaborated in the previous section, many pesantren institutions, including the one in this study, rely heavily on learners' interaction inside and outside the classroom as a medium for L2 learning and a previous study has reported non-target-like forms by the learners in such an institution (Aziez, 2016), which is common in peer interaction (Sato, 2015; Loewen & Sato, 2018).

To fill this interactional research gap, the present study is longitudinal and scrutinizes the development of English learners in an Islamic boarding school in Indonesia over time using a dynamic usage based (DUB) perspective (see Langacker, 2009; Verspoor & Behrens, 2011; Verspoor, Schmid & Xu, 2012; Roehr-Brackin, 2015). This perspective holds that the development of L2 learners depends on the learners' exposure to and experience with the L2. Larsen-Freeman (1976) argued that frequency of input has a significant role in the process of acquisition. From a DUB perspective, initial

conditions of the learners are very important too and, therefore, learners are expected to have different individual trajectories in their development. Learners' personal and linguistic background such as L1, scholastic aptitude, motivation, etc. are assumed to serve as predictor variables, which interact in complex manners and determine the acquisition of the L2. Furthermore, it is also believed that sub-systems of any organism are in some way interconnected and affect each other continuously in the development process (de Bot, Lowie, & Verspoor, 2007; van Geert, 1991). With regard to L2 development, the DUB approach looks beyond the division between linguistic features such as morphology and syntax. Morphology, lexicon, collocations, and sentence constructions are perceived as constructions in a linguistic continuum, which continuously interact as the L2 develop. Therefore, in investigating L2 development, one should examine as many sub-systems as possible to see not only how each sub-system develops but also how they interact.

With regard to English learning in a pesantren, it is also intriguing to see how the English of learners in such institution develop, particularly with the immersive nature of English language learning in many pesantrens in Indonesia including the one in this study. Hammerly (1991) criticized such immersive approach and argues that although the learners were able to attain a high level of communicative proficiency (fluency) with this approach, they tend to fail in terms of linguistic accuracy, leading to 'an error-laden classroom pidgin' since the learners are 'under pressure to communicate and are encouraged to do so regardless of grammar' (1991: 5). Therefore, it is suspected that the extensive interaction of the learners in English outside the classroom in a pesantren with a minimum exposure to the target language and correction from teachers and/or other learners (Bin Tahir, 2016) may lead to a pidginized form of English.

The resemblance of learners' language and pidgin languages has been pointed out by many linguists. Richards (1974) closely compared pidgin languages and second language acquisition (SLA). He argued that both codes can be described 'as an IL arising as a medium of communication between speakers of different languages, characterized by grammatical structure and lexical content originating in differing sources, by unintelligibility to speakers of the source languages and by stability' (Richards, 1974: 77). Schumann's (1978) famous study on Alberto, a Spanish learner of English who immigrated to the United States as an adult, indicated further the similarities between the structures of pidgin languages and the language of L2 learners. Many believe that

comparing learners' language with pidgin languages may shed some light on the emergence and development of pidgin and creole languages.

The emergence and development of pidgin and creole languages usually involve extreme case of language in contact such as slavery trades in the past, causing accelerated linguistic change (Levebfre, 2004), and it is near impossible to observe nowadays. It has been argued however that, in the beginning, second language acquisition plays a crucial part in shaping the languages followed by first language acquisition in its development when the speakers' children were exposed to the languages. One failed attempt to imitate such an extreme condition was done by Derek Bickerton and Talmy Givón in 1979, who proposed an experiment in which people speaking mutually unintelligible languages are taught approximately 200 words of English and then placed on an uninhabited island for a year where they would communicate using only the English lexicon while performing agricultural activities (as cited in Master, Schumann, & Sokolik, 1989). Their research proposal was obviously rejected due to the potential dangers to the participants of the study. Now, the previously described language learning situation in a pesantren may be able to provide this elusive context.

To sum up, we will trace the L2 use and development of students at a pesantren for one academic year, and take two cohorts, first year and second year students to simulate a two-year longitudinal study. There are four main questions that the present study attempts to answer. The questions are the following.

1) How do the learners at a pesantren interact in oral production and to what extent do the interactional features (trigger, corrective feedback, and modified output) occur in the learners' interaction? We will look at first-year and second year students and see if they differ in terms of interactional features.
2) What individual differences in terms of gender, motivation, scholastic aptitude in terms of class rank, age of acquisition of English, and initial writing proficiency predict the English writing development of the learners in the pesantren? Again, we will see if first-year students differ from second-year students.
3) To what extent do the learners' texts change overtime from a Dynamic Usage Based perspective in terms of holistic scores? Do the learners show variability over time or do they not? If not, to what extent do the learners show signs of stagnation in their L2 development? Do the learners show variation among each other? Again, we will see if first-year students differ from second-year students.

4) To what extent do we find elements of pidginization in the learners' L2? Which pidginization features are the most dominant and are there differences among first- and second-year learners?

CHAPTER 2
METHODS

The current chapter presents the methodological issues of the dissertation. At a *pesantren* (a boarding school in Indonesia), in which second language learning is assumed to benefit greatly from peer interaction several studies were conducted to explore the learners' L2 English development. Study 1 examines the peer interaction among students, focussing on several interactional features including trigger, negative feedback and modified output, which are believed to be important features for L2 learning. Study 2 explores the individual differences such as language background, motivation and scholastic aptitude that may affect L2 writing development over time. Study 3 explores L2 development over time and examines degrees of variability and stagnation. Finally, Study 4 explores written data for signs of pidginization.

This chapter deals with all the methods, procedures and analyses within the greater study and will deal with the separate studies where needed. Section 2.1 presents the research design. Section 2.2 describes the greater context in which this study took place. Section 2.3 presents relevant information of the participants. Section 2.4 discusses (a) how the data were gathered, (b) what instruments were used to measure different variables, and how the variables were operationalized. Section 2.5 present the analyses for each study and Section 2.6 summarizes each study and its specific research questions.

2.1. Research design

The current research was in essence a mixed method study. It includes both descriptive and statistical data in a longitudinal study aiming at exploring the practice of extensive peer-interaction and its impact on the learners' English development. The descriptive approach was mainly used in describing the interactional features and the pidginization features produced by the learners. The longitudinal approach was used in exploring how the learners English and pidginized forms (P-Forms) develop. Such dense longitudinal collection of data is particularly important in a second language development (SLD) study so we can gain better insight into the process of SLD (Verspoor, et al., 2008; van Dijk, et al., 2011). The research was conducted for one academic year at a pesantren institution in Indonesia. To simulate a two-year longitudinal study, two groups were involved in this study. The first group consists of first year students and the second group consists of second year students. The study began in their first week at the pesantren. The

participants in this study were involved in various regular programs designed by the school, which require them to participate in interactional activities in the L2. The learners were required to communicate in the L2 in their daily activities outside the classroom. This non-traditional approach allows the learners to engage in extensive peer-interaction as opposed to the brief amount of interactional treatment provided in the previous studies (M=30min) (Mackey & Goo, 2007).

There were several sources of data used in this study. The first source of data is the learners' conversations, which were analysed for their interactional features. Several surveys and questionnaires were used for to explore individual differences: a Language History Questionnaire (LHQ), motivation survey, and academic rankings were used in examining the predictors of the learners' English development. Development was operationalized as gains in holistic scores in the first few and last few writings. Texts written by the learners (i.e., 18 sessions in total, done every other week) in their English classes as part of learning process during one academic year were used to trace the learners' English development.

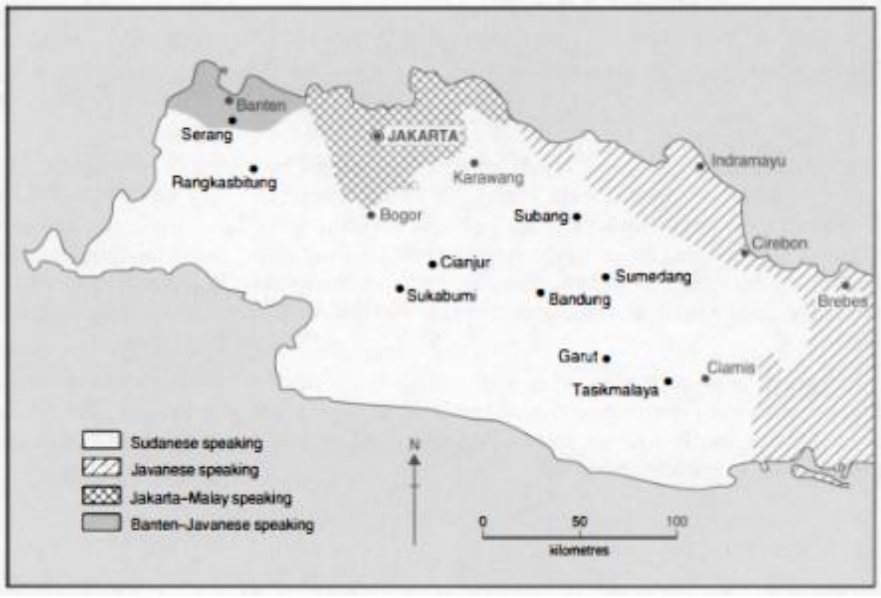
The first analysis was performed on the learners' interactions. Their interactions were examined for the extent of trigger, negative feedback and modified output, which were believed to be important features for L2 learning. In the second analysis, linguistic and non-linguistic backgrounds of the learners gathered through the questionnaire, motivation essay, and class rank were used to determine whether any of them correlate to the gains of the learners. To measure the gains, a pre-post approach was employed. For the pre- and post-scores, the average scores of the first three writings (pre) and the average scores of the last three writings (post) were used. Then, to get a better observation of the learners' progress, the average scores of the middle three writings (mid) were also used in the analysis. The third analysis was carried out on the holistic development of the learners' English based on their writings over time, looking for developmental patterns in terms of variability and stagnation. In the final analysis, samples of learners writing were analyzed for the extent of the features of pidginization.

2.2. Research context

This research was conducted in a pesantren in Tasikmalaya, a city in the province of West Java, Indonesia. Sundanese is used widely among the population of this area with the number of its speakers representing approximately 15% of the country's population

(Anderson, 1997). The following map (Figure 1) shows the linguistic map of the western part of Java where Sundanese is the dominant language in the region.

Figure 1. Linguistic map of the relevant western part of Java island (Anderson (1997), after R.R. Hardjadibrata (1997), Sundanese: a syntactical analysis, p.2. PL, D-65.)



Harsojo (1983:300-301 as cited in Anderson, 1997) summarized the macro level of the situation of Sundanese language in the area:

> Nowadays Sundanese is used widely among the population of West Java. In villages, the language of instruction is Sundanese, whereas, in towns, Sundanese is utilized primarily in the family circle, in conversation among friends and intimate acquaintances, and also in public and official places between people who are aware they both know Sundanese. With regard to language refinement, it is often said, that pure and refined Sundanese is to be found in the area of Priangan, that is, in the regencies of Ciamis, Tasikmalaya, Garut, Bandung, Sumedang, Sukabumi and Cianjur. Even now, the Cianjur dialect is still considered the most refined Sundanese. Considered less refined is the Sundanese near the north coast of Java, for example, that spoken in Banten, Karawang, Bogor and Cirebon.

Sundanese is one of more than 700 languages in Indonesia (Eberhard, Gary, & Charles, 2020) and with the multilingual nature of the country, the forefathers of the country saw the need of a unifying language. Indonesian, a standardized form of Malay, was then chosen as the official language of Indonesia, which serves as the lingua franca

of the archipelago (Sneddon, 2003). Indonesian is used as the language of administration, education, commerce, and the media. Consequently, almost all Indonesians speak the language to varying degrees of proficiency and since they already speak other regional languages as their L1, plurilingualism is the norm in the country (Zein, 2020). Although most Indonesians have a regional language as their L1, with the extensive use of Indonesian especially as the language of education, the number of Indonesian as L1 speakers is growing continuously. Based on the government's 2010 census, more than 40 million people in the country speak Indonesian as their L1 (Badan Pusat Statistik, 2010). According to the Indonesian government's regulation, Indonesian students have to learn both Indonesian and a regional language, particularly the language of where their school is located. Besides learning these two languages, Indonesians also have to learn at least one foreign language in almost all school levels. This regulation applies both to the public school system and the pesantren system, including the pesantren where this study was conducted.

Before going further into the language learning situation in this pesantren, it is important to understand the education level of the pesantren. Although many pesantren institutions begin with the elementary school level, the pesantren institution where the current study was carried out consists of two levels of education—the junior high school level (grade 7-9) and senior high school level (grade 10-12). This study will focus only on the first two years (grade 7 and 8). Each grade of the junior high school level consists of four learning groups, two female groups and two male groups. The number of students in higher grades usually have fewer students per group because many students move to public schools in the process because they cannot handle the high intensity of the learning process in the pesantren. This is understandable since the pesantren has almost three times more school subjects than public schools in general. Moreover, being away from their parents makes it harder for these young students. Usually, by the end of the junior high school level, less than 50% of the students remain and continue to the senior high school level.

In the pesantren, there are two compulsory foreign languages that the students have to learn i.e., Arabic and English. However, this dissertation will focus only on the latter. The school adopted its foreign language learning programs from the pesantren of Gontor (see van Bruinessen, 2006) since many of the teachers graduated from that institution. Like Gontor, the school also obliges its students to use Arabic and English in their everyday communication. The students have to use English and Arabic alternately

every week. The school has two curriculums on which their learning programs are based. According to the national curriculum, the students receive two lesson hours (160 minutes) of English. In this pesantren, the students also get another two lesson hours (160 minutes) of English reading class, which is part of the school curriculum.

There are also many additional activities in which students get their exposure of English language. Besides the classroom activities, the students get a daily vocabulary session called *mufradat* for fifteen minutes. The words of the day are taken from a book which they obtained from another pesantren (see Appendix E for sample). In this session, they get two to three English words from an appointed senior student from grade 10 (senior high level). After some pronunciation drills, they are asked to make English sentences using the given words (Figure 2).

Figure 2. A *mufradat* or vocabulary session

These sessions are carried out six times a week during English weeks. For productive skills practices, the students' main public speaking sessions, which are called *muhadharah,* are conducted every Tuesday and Thursday for about one-hour period in each session. In this session, students have to give a speech on religious topics in Indonesian, Arabic, and English (Figure 3). There are usually around 20 students in a mixed group of different grades. However, a student commonly performs as a speaker once in each language in one semester period. Most of the time, the students participate as audience.

Figure 3. *A* muhadharah *or public speaking practice session*

On Tuesday and Friday mornings of the English week, they do half-an-hour English conversation practice called *muhadasah*. It is usually done outdoor in a field where the students stand in two lines facing one another. They are then given a topic by a teacher or an appointed senior student. Then, they start the conversation while being supervised by the teacher or the senior student. The supervision mainly focuses on maintaining the flow of conversation—making sure no one stops speaking. Students sometimes ask the supervisors for the meaning of some Indonesian or Sundanese words in English but from observation we know that very little correction was made by the supervisors when the students mispronounced a word or made grammatical errors. Figure 4 below shows the session and a senior student with blue outfit can be seen walking between the lines of students.

Figure 4. A *muhadatsah* or conversation session

The proficiency goal of language learning in the *pesantren* is lexically based with the learners being expected to have acquired 4000 words of Arabic and English when

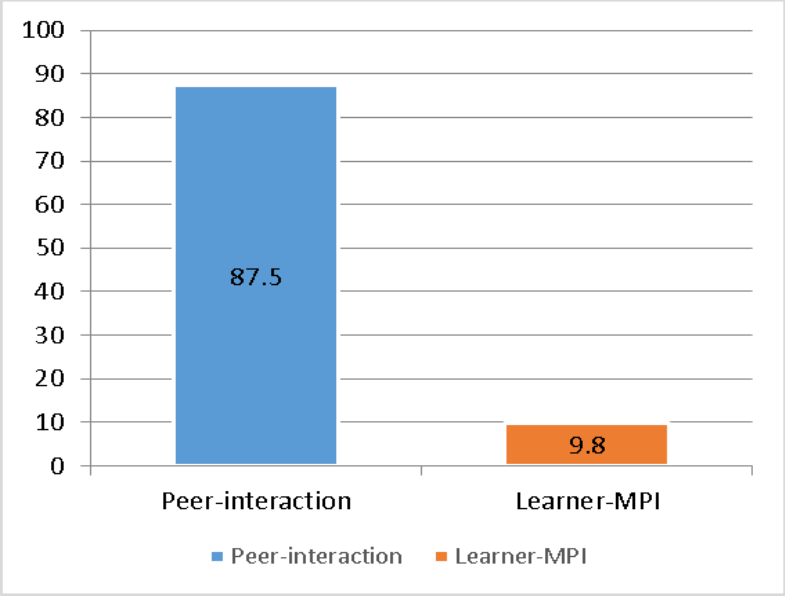
they graduate. Grammar was taught but not to the same extent as vocabulary. When the students are caught using an Indonesian or Sundanese word, for example, they are given a list of English or Arabic words to be memorized, depending on what week it was, as a form of punishment. On the next day, they had to come to one of the appointed senior students to get tested on their memorization of the given words. Similar practice has also been reported by Jihad (2011) in other pesantren institutions. Therefore, the aforementioned *mufradat* sessions were also seen to be very important in building the learners' vocabulary mastery. In the first few sessions after their enrolment, students were given English and Arabic words of the things found around the pesantren to help them communicate and get around the school complex. This is very important for them since they were obliged to speak in those two languages after only three months in the pesantren. To help learners with the words they needed, the school put up some lists of words, which were usually related to where the lists were displayed. Figure 5 below shows some examples of the lists of vocabulary displayed in the sport yard, school clinic, and kitchen.

Figure 5. Vocabulary lists displayed in some areas of the school

The second goal of language learning in the pesantren is fluency and the school relies on peer interaction to reach this goal. It is observable that peer interaction has a significant portion in the language learning process in the pesantren. This was made possible by the fact that the students live inside the school complex. Compared to the students in public schools, which commonly have only 160 minutes of English class, the students in the pesantren have virtually indefinite exposure to English, especially during

English weeks. However, as discussed in the previous section, there is some doubt about the quality of input and feedback that can be provided by fellow learners (e.g., Adams, Nuevo & Egi, 2011) compared to more proficient interlocutors such as English teachers. From the observation of the students' activities during the English week, Figure 6 compares the number of hours per week in which possible interaction can occur between the students with more proficient interlocutors (MPI) (i.e., teachers and appointed seniors) and students with their peers.

Figure 6. Comparison of possible interaction time between two dyads (hour/week)



The graph clearly shows a great difference in the amount of time in which interaction can occur between the two dyads. Although some research has pointed out the psychological and L2 learning advantages that peer interaction can provide in classroom settings, there has been no study conducted to measure the impact of such extensive learner-learner interaction on L2 development particularly in the naturalistic context such as this pesantren in Indonesia.

2.3. Participants

The participants in this study were first year (grade 7 junior high school level, hereinafter Group 1) and second year (grade 8 junior high school level, hereinafter Group 2) students of junior high level at the pesantren. In the beginning, 126 first year students and 85 second year students participated in this study. However, the following exclusion criteria were applied to ensure the validity and reliability of the present study:

a. Missing questionnaire
b. Missing motivation essay
c. Missing two or more sessions in a row
d. Missing more than once in the first three, middle three, and final three sessions
e. Absence in 5 sessions or more in total
f. Dropping out of school during the research period

Based on the exclusion criteria, 44 first year students and 29 second year students were eventually excluded. In the end, the data from 82 first year and 56 second year students were included in the final analysis. Table 1 shows the number of learners based on their gender.

Table 1. Number of learners based on gender

Gender	Group 1 (n)	Group 2 (n)
Male	39	25
Female	43	31
Total (n)	82	56

Table 2.1 above shows that generally, there are more female learners than male learners in both groups. The age of learners from Group 1 ranged from11 to 13 with an average age of 12.2, while the learners' age from Group 2 ranged from 12 to 14 with an average age of 13.1. The average age of the male learners is slightly higher than of the female learners as seen in Table 2 below.

Table 2. Learners' average age based on gender

Gender	Group 1	Group 2
Male	12.38	13.28
Female	11.95	12.94

Data collected by means of the Language History Questionnaire (LHQ) version 2.0 which had been translated into Indonesian (see Appendix A) showed the information about the participants themselves and their linguistic backgrounds. The results from the LHQ shows that almost all learners were multilingual with Sundanese as the L1 for most of them. A more detailed results of the LHQ will be presented in the results chapter.

2.4. Procedures

After finding the potential pesantren for the study, the researcher went to the pesantren personally one week before the commencement of the learning activities in the new academic year. The researcher had a meeting with the principal of the pesantren to ask for informal permission as well as to discuss the study plan. After the meeting, an official letter requesting permission to conduct a study at the pesantren was then submitted to the principal office. After getting the permission, the principal then appointed the English teacher who would help the researcher in gathering the necessary data. The principal also provided the researcher with accommodation for the first week within the pesantren complex. This allowed the researcher to conduct the first important steps of the research as well as to get a thorough observation of the learners' activities at the pesantren. It should be noted that, at the beginning of the study, after acquiring permission from the school principal, an informed consent (see Appendix B) was given to and signed by participants and the parents or guardians of the participants for ethical conduct of this study. The consent form consisted of a description of the study, the research processes, and statement of confidentiality of the data collected during the study.

During this period, observation was carried out for a full week during an English week to understand the extent of the students' interaction in English from the time they wake up until they go to bed. Several scheduled learning activities were observed including their English classes, reading classes, conversation practice sessions, vocabulary sessions, and public speaking practice sessions. Moreover, their daily activities outside classroom were also observed to see when and where the students usually interact. In addition, the documents on curriculum, schedules, school rules, and academic and nonacademic activities were collected from the school administration with the permission from the school principal. Interviews were also carried out with the principal and the teachers of the school to confirm and clarify particularly the information from the mentioned documents.

The details in which each study was carried out will be elaborated in the following sub-sections.

2.4.1. Learners' interaction

At the beginning of the study, 8 pairs of learners were picked randomly (4 pairs from 1st year and 4 pairs from 2nd year). Since the learners had their own activities during leisure time, the teacher asked any students who were free at the time to come to the teacher room to perform the interaction task. They were asked to have a conversation with

their partner on their daily activities at the pesantren. The learners were left alone in the room without the presence of the teacher and the researcher to avoid nervousness. A voice recorder was set in the room to record the interactions.

The recorded learners' interaction was transcribed by the researcher. The transcript was then examined by the researcher and one other rater for the frequency of the interactional features namely trigger, negative feedback and modified output which are believed to be important features for L2 learning. Trigger is simply learner's non-target-like utterances. Negative feedback and modified output were described as follow:

- Negative feedback

Recast: A learner's more target-like reformulation of his/her interlocutors' non-target-like utterance. The reformulation of the interlocutor's utterance could be partial or complete.

Example:

Student A: The English lesson is one hours.

Student B: One hour.

Clarification request: A learner's attempt to elicit information from his/her interlocutor using any form of request for clarification, such as what, pardon, huh, etc. Clarification requests in English, Sundanese, or Indonesian were included.

Example:

Student A: What your favorite food?

Student B: Eat, eh?

Student A: Food!

Explicit correction: A learner's explicit statement that the interlocutor's utterance was incorrect. The correction may include metalinguistic explanation or explanation in Sundanese or Indonesian.

Example:

Student A: I eat yesterday.

Student B: No, it should be past 'ate'.

- Modified output

Modified output is a learner's reformulation of his/her previous non-target-like utterance which results in a more accurate form. Modified output can be a response to an interlocutor's feedback as well as self-initiated.

Example (in response to feedback):

Student A: How many hours in a day?

Student B: *Dua*

Student A: Two hours?

Student B: Two hours.

Example (self-initiated):

Student A: Does you have uh… Do you have English lesson?

Besides the frequency of these interactional features, several other aspects were also examined including the number of turn takings, target-like utterances and non-target like utterances.

2.4.2. Individual differences

On the first day of the visit, the researcher met with the English teacher to set a common understanding of the practical and theoretical ground of the study. A plan was then set for that week. The first three days were spent on administering the LHQ to all the participants of the study. In each session, after filling the LHQ, learners were asked to write a short motivation essay which states why they study there and whether it was on their own initiative, their parents or family, or both. The researcher and the teacher were present during each session. After completing the LHQ and the essay, the participants then collected and handed them to the teacher or the researcher. Whereas the LHQ and the motivation essay were collected at the beginning of the study, the class rank data was collected after the first semester ended, as the academic reports were only available at this time. The procedures of each instrument in this part of the study will be elaborated individually below.

2.4.2.1. The Language History Questionnaire

To obtain linguistic background of the participants, an Indonesian translated version of Language History Questionnaire (LHQ) 2.0 (see Appendix A) were administered. LHQ is a widely used tool for assessing the linguistic background of bilinguals or L2 learners and for generating self-reported proficiency in multiple languages (Li et al., 2014). The questionnaire is available for free on the website of Brain, Language, and Computation Laboratory (BLCLAB), The Hongkong Polytechnic University.

2.4.2.2. Motivation

The LHQ was a long questionnaire and to avoid boredom, motivation was measured with free response data. To measure learners' motivation, the learners were asked to write a reflection in about 100-200 words (see Appendix C for sample), in their L1, on their motivation to enrol in the school. They were instructed to write about why they study there and whether it was on their own initiative, their parents or family, or both. Then, the learners' reflections on their motivation were scored based on Self-Determination Theory (SDT), a motivation theory focusing on individuals' motivation-related qualities and motives that affect their behavior (Utvær & Haugan, 2016). It emphasizes the integration and regulation of personal motives within the self. However, as one's motivation cannot be separated from his or her interaction with the environment, SDT also considers how the self internalizes motivation under the influence of the inseparable social contexts (Deci & Ryan, 2014). Autonomous types of motivation are seen as high-quality motivation. In contrary, controlled motivation is thought to be of low quality (Deci & Ryan, 2000, 2008). Utvær and Haugan (2016) presented the internalization continuum as well as the various types of motivation that they created by adapting a scale from Deci and Ryan (2008), and Ryan and Deci (2009) (Figure 7).

Figure 7. The types of motivation and regulation within SDT

<table>
<tr><td>Amotivation</td><td colspan="4">Extrinsic motivation</td><td>Intrinsic motivation</td></tr>
<tr><td>Non Regulation</td><td>External Regulation</td><td>Introjected Regulation</td><td>Identified Regulation</td><td>Integrated Regulation</td><td>Intrinsic Regulation</td></tr>
<tr><td>Least self-determinated</td><td colspan="4">←——————→</td><td>Most self-determinated</td></tr>
<tr><td>Amotivation = lack of motivation</td><td colspan="2">Controlled motivation = low-quality motivation</td><td colspan="3">Autonomous motivation = high-quality motivation</td></tr>
</table>

Instead of perceiving internalization and types of regulation as either intrinsic or extrinsic, SDT have shifted that conception to one of autonomous and controlled types of motivation. Based on the above continuum, the following categories (Table 3) were created and used in rating the participants' motivation. The rating of the learners' motivation was done by two raters.

Table 3. Motivation scoring category

Type of motivation	Description	Code
Amotivation and controlled motivation	Learners' show lack of motivation or learners' motivation comes from outside (e.g., parents or family members) and little to no indication of internalization.	1
Autonomous motivation	Learners' motivation comes from own self or outside (e.g., parents or family members) but shows significant indications of internalization.	2

2.4.2.3. Scholastic aptitude

To estimate language aptitude, a standard language aptitude test should be used. However, in the current study with so much data to be collected, we did not have the time to administer such a test to so many students. Therefore, in line with Verspoor, de Bot and Xu (2015), who found a significant correlation between scholastic aptitude and English proficiency gains, we chose to operationalize aptitude in terms of scholastic aptitude. Students' academic report showing the class rank was used as a measurement of the learners' scholastic aptitude. The documents of the learners' academic report were provided by the school with the permission of the principal and the academic counsellor. They are categorized based on the following order; learners who belong to the top 20% in their class were coded 1 and so on. The categories are presented in Table 4 below.

Table 4. Class rank categories

Category	Rank in their class
1	≤20%
2	21%-40%
3	41%-60%
4	61%-80%
5	81%-100%

2.4.3. L2 development

The dependent variable in this study, EFL development, was operationalized as English writing development. This study uses free writings in the form of narratives as a means to observe the English language development of the learners as well as the

pidginization process. The selection of English writing development has been driven by theoretical and practical considerations.

Theoretically speaking, assessing language learners' writing has been used as one of the ways to measure their general proficiency in L2. Verspoor, Schmid and Xu (2012) argued that written texts can exhibit learners' active language use, rather than passive knowledge of L2, in all its facets, such as the use of vocabulary, idioms, verb tenses, sentence construction, and errors. They also added that writing can mirror learners' target language proficiency better than speaking since it gives space for more reflection and therefore can dig more into what the learners can do with the target language. Moreover, as an additional benefit for researchers, writing data is considerably easier to collect and assess than spoken data. Therefore, the development of language learners' writings is viewed as a representation of the development of their English proficiency.

As for practical consideration, written data are considered to be easier to collect and process than spoken data. Moreover, the exercise of free writing has also been used successfully in several other studies on language proficiency (e.g., Verspoor, Schmid & Xu, 2012; Hong, 2013; Verspoor & Smiskova, 2012; Irshad, 2015). It is also proven to be suitable for beginners (Crowhurst & Piché, 1979). An analysis of the first sample of the learners' writings also shows similar features with their oral form of English as presented in the previous study done in the same context (Aziez, 2016). Therefore, it was considered feasible to choose learners' writing as measure of their English development.

The writing sessions were carried out over a one-academic-year period. The learners participated in 18 writing sessions in total, conducted once every other week in their English classes. Every writing session lasted approximately 20 minutes. The first writing session was conducted with the presence of the researcher and the English teachers. Since the researcher could not be present for the whole year at the school, the rest of the sessions were conducted by the teachers only with constant communication with the researcher. After each writing session, the teachers uploaded the files to a cloud file sharing service.

In the first session, the teacher emphasized to the learners that their writings would not affect their English subject grade at school. The learners were also asked to write as many words as possible in their writing and told that their writings would be graded anonymously, which means no attention was paid to individual errors at this stage. The writing sessions were also included as part of their English classes so that they can practice their English writing skills. This emphasis was important in order to minimize

the detrimental effect of anxiety on the learners' language production as shown in several studies (e.g., Horwitz et al, 1986; MacIntyre, 1995; Dornyei, 2005). Moreover, learners were not allowed to consult any dictionary or get any help from their peers or teachers. These restrictions were implemented to make sure that the writings reflect the learners' actual English proficiency.

Before the first session, the researcher and the teachers agreed on the following topics (Table 5) for the writing sessions based on the topics provided by the curriculum or based on recent events. The topics were applied to both the first and second groups. At first, we considered choosing the same topic (i.e., daily diary/my activity today). However, the teacher suggested that this would be boring for the learners since they generally have a strict daily plan set by the school. This would also allow learners to copy from their friends since they have common daily activities.

Therefore, we chose different topics based on several considerations including the linguistic features commonly used for the topics. Although the topics may look different, they are all narrative and descriptive in nature and share similar linguistic features. They are also repeated throughout the sessions. Moreover, the topics selected are closely related to the learners' life at the school to make sure that the learners are familiar with the topics which is also important in writing assessment (Schoonen, 2012). This would allow learners to have sufficient background knowledge on the topics as well as to repeat the same linguistic features which could be traced overtime. Schoonen (2005; 2012) reported that linguistic features have been reported to be one of the aspects that are least affected by the tasks. He also added that when writing texts are scored holistically as in our study, it is more generalizable than analytic scores (Schoonen, 2005; 2012).

Table 5. List of topics

Session	Topics
1	About myself
2	Celebrating Independence Day
3	My favorite food
4	My favorite movie
5	My favorite place in the school
6	My happy experience
7	My best friend
8	My hobby
9	My last holiday
10	My favorite lesson and teacher
11	My family
12	My free time in the school
13	My favorite game
14	My classmates
15	My hometown
16	Celebrating *Idul Adha*
17	Celebrating *Idul Fitri*
18	My dream

All the learners' interaction recordings and handwritings were transcribed by the researcher into a word file to make it easier to analyze. The researcher chose approximately 25% from the transcriptions and asked a rater to check the accuracy. The inter-rater reliability was calculated by transcription agreement rate which was 97%.

In total, there were more than 1000 pieces of writings from all sessions. However, writings from excluded participants were ultimately not transcribed. All proper nouns in the writings were coded (name) to ensure the confidentiality of the data. Unclear handwritten words were consulted with the rater and if no conclusion was made, the word(s) were coded as [unclear].

To assess the general proficiency of the learners, their writings were holistically scored. Ortega indicated that studies that used holistic ratings have resulted in more homogeneous observations as reflected in smaller standard variations and narrower ranges for the measures she has investigated than those that assess proficiency in terms of naturally occurring classes or groups (2003: 502). The procedure of assessing the texts was controlled carefully to maintain high inter-rater reliability. In developing the scoring criteria, the steps used in Verspoor et al. (2012) were adopted. A group of five raters developed the scoring criteria as follows: Every rater was provided with six samples,

which they assessed in order to determine who they believed to be the strongest and the weakest in English. The raters then discussed these orders. A variety of features that are closely associated with general CAF indicators arose from the discussions between the raters: text length, sentence length, sentence complexity, use of different types of clauses, use of tense, vocabulary range, use of L1, use of idiomatic language, and accuracy. Some sample texts were difficult to score, which resulted in lengthy discussion among raters before agreement was made. After a rating agreement was reached, the texts were provisionally graded into different levels of proficiency. With this procedure, the raters worked with several samples until they achieved six levels of proficiency (0–5). From the discussion, the raters also agreed on half scores (e.g., 1.5, 2.5, etc.) for texts that have features of two different scores. The criteria for the holistic scoring are presented in Table 6 below.

Table 6. Holistic Scoring Criteria

Score	**Descriptor**
0	No attempt made despite being present.
1	Short text (less than 50 words). Very short sentences. Possible mix with L1. Phonetically spelled. First attempt at L2.
2	Short text (around 50 words). Very short sentences. Possible mix with L1. Mostly present tense. Very simple vocabulary.
3	Longer text (over 80 words). Mostly English but literally translated from L1 (confuses reader). Mostly main clauses and attempt at dependent clause. Simple vocabulary. No chunks. Attempt at use of other tenses. Not coherent. Jumps from one topic to the other.
3	Longer text (over 80 words). Mostly English. But not all understandable (lots of misspellings or words left out). Mostly main clauses. An attempt at dependent clauses. Simple vocabulary. No TL chunks. Use of present tense. Coherent story.
4	Longer text (over 100 words). Mostly English. Mostly main clauses. Simple vocabulary. No chunks. Use of present tense. Using past tense, not quite consistently yet. Coherent story.
5	Longer text (over 100 words). Mostly English. Use of dependent clauses. Some less frequent words. Some chunks (but also some not TL). Use of other tenses. Coherent story.

After the above criteria for assessing the texts were set, the researcher ranked the rest of the writings using the criteria with the help of one rater. Another rater was consulted when there were discrepancies in scoring the text.

2.4.4. Pidginization

As mentioned at the beginning of the chapter, samples of learners' writings were also analysed for signs of pidginization. For this purpose, writings from 10 learners from each group (20 in total) were scrutinized for pidginization features. In choosing the samples, correlation analysis was used to decide which sets of learners' writings were used. The learners whose writing scores had the strongest positive correlation with the average scores of the group were chosen.

The chosen texts were examined for the following characteristics of pidginization as suggested by Andersen (1981, as cited in McLaughlin, 1987), as well as Schumann (as cited in Johnson & Johnson, 1999):

(1) A basic pidgin negation, e.g. "I no see".
(2) Lack of inversion in questions, e.g. "Where he is?".
(3) Lack of auxiliaries, e.g., "She crying";
(4) Lack of possessive inflection, e.g., "The king food".
(5) Lack inflectional morphology, e.g., "Yesterday, I eat noodle"; subject pronouns, e.g., "No have holidays".
(6) Use of L1 words, e.g., "my father *bangga".*
(7) L1 based forms and construction, e.g., "food nice"; "I life in this boarding".

There are two characteristics mentioned in the references that were excluded from the examination since they did not appear in the linguistic context of the learners in this study. The first characteristic is reliance on word order rather than inflections for expressing grammatical relations. The second characteristic that was excluded is sporadic merging of pre-verbal markers which come from lexical verbs promoted to auxiliary status. In addition, we counted simplifications and transfers of the source language, which are also typical of the language of a learner, to see if they remain at the end of the school year and if they did, we considered them pidginization features.

2.5. Analyses

2.5.1. Learners' interaction

The researcher and the rater examined eight conversation samples for the interactional features namely turn takings, target-like utterances, non-target like utterances (also called trigger), negative feedback, modified output. Then, using simple

percentage calculation for inter-rater reliability, 100% agreement was reached for the negative feedback categories and 97% for modified output.

The frequency of the interactional features in the first-year students' and the second-year students' interactions were also compared to see whether there were any differences between the two groups. However, as they were so low, no statistical analysis was performed.

2.5.2. Individual differences

The analyses for this study were carried out on the writing score as the outcome variable. First an independent t-test was used to see whether the difference in the writing scores of Group 1 and 2 is significant. Due to the nature of the variables (continuous, categorical, binary), two types of analyses were used to explore the relationship between the independent variables (gender, class rank, motivation, age of acquisition, initial writing score) and the outcome. All analyses were carried out for each group (1 and 2) separately. A one-Way between subjects ANCOVA was used with class rank (fixed factor, ordinal variable), initial writing proficiency (covariate, continuous variable) and the outcome writing scores (dependent, continuous variable). A linear regression analyses was also run for the two groups separately with gender, initial writing proficiency, motivation and age of acquisition as predictors and the writing score as the outcome.

2.5.3. L2 development

The writing scores were processed using the analytical software SPSS 22.0. To get an overview of learners' development, descriptive analyses were first carried out. The scores from the participants in every session were averaged and compared based on groups and gender to see the overall development of the groups. Then, to determine if there was any significant progress of the learners' writing scores, the pre-post approach was employed. For the pre-and post-scores, the average scores of the first three writings (pre) and the average scores of the last three writings (post) were used. Then, to get a better observation of the learners' progress, the average scores of the middle three writings (mid) were also used in the analysis. By averaging these three scores, we hoped to avoid the effects the different topics and the missing data on the overall scores. To test the normality of the distribution of the data, *Kolmogorov-Smirnov* test was performed. Then, *Levene's test* was also carried out to test the homogeneity of the data. When the data were normally distributed and homogenous, then ANOVA and independent t-test

were performed. In contrast, when the data were not normally distributed and not homogenous, the data were analysed using non-parametric tests namely *Mann-Whitney* and *Kruskal-Wallis* tests.

In variability analyses from a CDST perspective, the trajectories of individual learners are inspected visually to see if scores go up or down rapidly from one session to the other or if there are major shifts. Visual inspection may be aided with min-max graphs or polynomial trend lines. (cf. Verspoor et al. 2011 for various techniques.) However, visual inspection indicated that none of the learners showed changes over time after the first few months. Also at the group level, the pre-post test showed that there is not much change over time. Therefore, no further variability tests were conducted. In the group analysis, variability for each learner was operationalized as the coefficient of variation (CoV), in line with Verspoor and de Bot (2021), but they also point out that this measure may be inadequate as it does not take time into account. They recommend that the Standard Deviation of Differences be used instead.

2.5.4. Pidginization

With the help of one other rater, the researcher examined in detail the writing samples of 20 learners for pidginization features. To select the samples, the learners' holistic scores in the writing tasks were correlated with the group average. The learners with the strongest correlation coefficient with their group's average scores were then selected. During the categorization process, discrepancies were discussed between the researcher and the rater until agreement was reached. Each pidginization feature was marked and counted. The percentage of the number of the features from the total number of words in each text was calculated. For the first step, a pre-post analysis of this ratio was carried out to see whether the learners improved in the sense that they produced fewer pidginization features overtime. For this step, we used the average of the ratio of session 2 and 3 for the pre-score and the average of session 17 and 18 for the post score. We did not use the first session since the topic of the first session is self-introduction, which apparently was very easy for the students and consisted mostly of well-memorized phrases. This was indicated with the fact that they produced significantly fewer pidginization features in this first session. Results of Group 1 to Group 2 were also compared to see whether there were any differences between the groups. We assumed that Group 1would improve overtime while Group 2 would remain stable. Finally, we

also counted the number of occurrences of each pidginization feature to see which features are more common in the learners' L2. We also compared the features found in Group 1 and Group 2.

2.6. Summary

This dissertation aims to explore the English language development of 138 young Indonesian learners in their first and second year at a pesantren, which emphasizes the use of peer interaction in their English learning process. Since these learners have little access to authentic English, we assumed that based on several theories there is a possibility that their reliance on peer-interaction for learning English could lead to pidginization. Altogether, there are four studies which shape this dissertation.

The first study explores the learners' interaction. It seeks to elucidate how the learners interact in oral production and to what extent the interactional features (corrective feedback, modified output, and self-initiated modified output) occur in the learners' interaction. Moreover, this study also tries to examine how the first-year students differ from the second-year students in terms of such interactional features. To do this, samples of learners' interaction from both groups were examined for the frequency of which the interactional features occur. Since the frequency of the interactional features turned out to be very low, no statistical analysis was done and the results will be presented in a descriptive manner.

The second study acknowledges the importance of the learners' individual differences in L2 development. Therefore, it attempts to answer which individual differences in terms of age, gender, motivation, scholastic aptitude in terms of class rank, self-reported language learning ability, age of acquisition of English, and writing proficiency predict the English writing development of the learners in the pesantren. Also, this study examines whether there were any differences between first-year students and the second-year students in this regard. In doing so, some statistical analyses were carried out including an independent t-test, a one-way between subjects ANCOVA, and a linear regression analysis.

The third study explores L2 development over time and examines degrees of variability and stagnation. Taking a Dynamic Usage Based perspective, it attempts to answer the question whether the learners' texts change overtime from in terms of holistic scores and whether the learners show variability over time and variation among each other. Also, it seeks to answer whether the learners stagnate at a particular point in time.

To determine if there was any significant progress of the learners' writing scores, the pre-post approach was employed. The data were tested whether they are normally distributed and homogenous. If they are normally distributed and homogenous, then ANOVA and an independent t-test were performed. In contrast, when the data were not normally distributed and not homogenous, the data were analysed using non-parametric tests namely *Mann-Whitney* and *Kruskal-Wallis* tests.

Finally, the fourth study explores the written data for signs of pidginization, especially to what extent we find features of pidginization in the learners' language. For this purpose, writings from 20 learners were used in the analysis. The ratio between the number of pidginization features and the total number of words in each text were calculated. The average ratio from the first two sessions was compared to the average ratio of the last two sessions to see whether the learners improve in the sense that they produce fewer pidginization features overtime as they were acquiring English. Results of Group 1 and Group 2 were compared to see whether there was any difference between the groups. Finally, we also counted the number of occurrences of each pidginization feature to see which features are more common in the learners' L2.

In the following chapter, the results of the four studies will be presented.

CHAPTER 3
RESULTS

This chapter presents the results of the analyses as elaborated in the previous chapter. Section 3.1 presents the results of the analyses on the peer interaction, particularly in terms of the interactional features including turn takings, trigger, negative feedback and modified output. Section 3.2 discusses the results of the second study on the effect of individual differences such as gender, language background, motivation and scholastic aptitude on the learners' L2 writing development. Beside these individual differences, this section also presents the analysis on the differences between the first group and the second group. Section 3.3 presents the analysis results on the learners' holistic L2 development over time. Then, section 3.4 presents the signs of pidginization that were found in the learners' L2. Lastly, section 3.5 summarizes the main findings of each study.

3.1. Learners' interaction

In this study, we seek to learn how the learners interact in oral production and to what extent the interactional features (i.e., turn takings, trigger, corrective feedback, and modified output) occur in the learners' interaction. Moreover, this part also attempts at finding out the difference between the first-year students and the second-year students in terms of the interactional features.

The interaction data gathered consist of transcriptions of audio recordings from 8 conversations between learners of the same group and gender. The length of peer interaction ranged from 3-13 minutes for each conversation with an average of 6 minutes. The total time of the learner-learner interaction in the data set is approximately 49 minutes. Table 7 shows the frequency of the interactional features in the data set.

Table 7. Frequency of interactional features

Interactional features	Type	Sum Group 1	Of which NTL	Sum Group 2	Of which NTL
Turn taking		107	54	286	124
Negative Feedback	Recast	0	0	5	2
	Clarification request	0	0	1	0
	Explicit correction	0	0	1	0
Modified output	As response to feedback	0	0	7	2
	Self-initiated	3	2	6	1

As seen in the table, there are more turn takings in Group 2 (n=286) than Group 1 (n=107). In the interactions between learners in Group 1, 54% of the learners' utterances are non-target like (NTL) while 43% of utterances in the interaction of Group 2 are NTL. However, it should be noted that the turn-takings also includes short answer such as 'yes' or 'maybe' and fillers such as 'uh', 'err', etc. None of the NTL utterances produced by the learners in Group 1 resulted in feedback. Thus, the only 3 instances of modified output were self-initiated and of which 2 were still NTL. In the interaction between learners in Group 2, 124 NTL utterances resulted in only 7 instances of negative feedbacks of which 2 were NTL. Of the 7 feedbacks, 5 were in the form of recast, 1 was a clarification request, and 1 was an explicit correction. All 7 instances were responded with modified output. However, 2 of which were still NTL. There were also 6 self-initiated modified output in the second group, one of which was NTL.

3.2. Individual differences

This particular study seeks to explore which individual differences in terms of age, gender, motivation, scholastic aptitude in terms of class rank, self-reported language learning ability, age of acquisition of English, and writing proficiency predict the English writing development of the learners in the pesantren. It also aims at finding out how first-year students differ from the second-year students.

Before presenting the statistical analyses of this study, the descriptive results of the instruments used to gather the data on the learners' individual differences namely language history questionnaire, motivation questionnaire, and documents on the class rank used as an indicator of scholastic aptitude. In the statistical analysis section, the result from the regression analysis will be presented.

3.2.1. Descriptive analysis

3.2.1.1. The Language History Questionnaire

To get information on the linguistic backgrounds of the learners, the Language History Questionnaire (LHQ) version 3.0 was administered. Table 8 below summarizes the results from the collected the LHQ submitted by the participants from both groups of participants.

Table 8. Results from LHQ

Variables	Categories	Group 1		Group 2	
		n	%	n	%
Multilingual/ Non-multilingual	Multilingual	79	94.3	54	96.4
	Non- Multilingual	3	3.7	2	3.6
Number of Languages	1	3	3.7	2	3.6
	2	44	53.7	27	48.2
	3	35	42.7	27	48.2
Native language	Indonesian	24	29.3	23	41.1
	Sundanese	58	70.7	33	58.9
	Javanese	0	0.0	0	0.0
	Other Languages	0	0.0	0	0.0
Self-report language learning ability	1	1	1.2	0	0.0
	2	1	1.2	0	0.0
	3	6	7.3	0	0.0
	4	34	41.5	19	33.9
	5	33	40.2	35	62.5
	6	7	8.5	2	3.6
	7	0	0.0	0	0.0
English Background	No English Background	33	40.2	15	26.8
	Little English Background	19	23.2	16	28.6
	Strong English Background	30	36.6	25	44.6
Frequency of code mixing	No mixing	20	24.4	10	17.9
	Low	46	56.1	41	73.2
	High	16	19.5	5	8.9
Comfortable in terms of writing	Indonesian	68	82.9	38	67.9
	Sundanese	14	17.1	15	26.8
	Javanese	0	0.0	0	0.0
	Other Languages	0	0.0	3	5.4
Comfortable in terms of speaking	Indonesian	28	34.1	21	37.5
	Sundanese	53	64.6	32	57.1
	Javanese	0	0.0	0	0.0
	Other Languages	1	1.2	3	5.4

The table above shows that most of the participants are multilingual (94.3% learners from Group 1 and 96.4% learners from Group 2) with 53.7 % of learners from Group 1 and 48.2% of learners from Group 2 having 2 acquired languages prior to their enrolment in the school and 42.7% learners from Group 1 and 48.2% learners from Group

2 having acquired 3 languages prior to their enrolment in the school. Only 3.7% and 3.6% of learners from Group 1 and Group 2 respectively were monolingual. It should be noted that the number of languages does not indicate the learners' level of proficiency in those languages. Sundanese is the L1 of the majority of the learners (70.7% of learners from Group 1 and 58.9% of learners from Group 2). Indonesian languages are the L1 of 29.3% learners from Group 1 and 41.1% learners from Group 2. In terms of self-report language learning ability, in the scale of 1-7, most of the learners scored themselves 4 and 5 (81.7% learners from Group 1 and 96.4% learners from Group 2). Regarding previous English exposure, 40.2% of learners from Group 1 and 26.8% learners from Group 2 had no previous English exposure prior of the enrolment to the school. The rest of the students have received, to some extent, some exposure of English during the elementary school. However, it should also be noted that this does not indicate their English proficiency. Most of the learners (75.6% learners from Group 1 and 82.1% learners from Group 2) mixed their codes to some extent with most of them claiming that they did not do it too often. For most of the learners, Indonesian is the language they are most comfortable with in terms of writing (82.9% learners from Group 1 and 67.9% learners from Group 2). This may be due to the fact that Indonesian is the primary instructional language in the country. Finally, Sundanese is the language that the learners claimed that they are comfortable with in terms of speaking (64.6% learners from Group 1 and 57.1% learners from Group 2). Interestingly, although the L1s of the learners were only Sundanese and Indonesian, a few learners filled other languages as the language they are comfortable with in terms of speaking and writing. When confirmed, they said that they moved to other regions in the country or abroad in their early age where their L1 is not the language spoken there.

Because almost all students were multilingual and the degree of multilingualism is so complex in terms of when the learners learned the languages and how well they speak and write them, it was impossible to operationalize them into one construct for the regression analysis. Therefore, we did not include it in further analyses.

3.2.1.2. Learners' motivation

Table 9 below shows the results from the learners' motivation questionnaire that have been categorized into amotivation/controlled motivation and autonomous motivation based on a scale developed based Utvær and Haugan's (2016) internalization continuum.

Table 9. Results from motivation essay

Variables	Categories	Group 1		Group 2	
		n	%	n	%
Motivation	Amotivation and controlled motivation	23	28	19	33.9
	Autonomous motivation	59	72	37	66.1

The table shows that most of the learners (59% learners from Group 1 and 66.1% learners from Group 2) have autonomous motivation and show indications of internalization. Less learners had low motivation or controlled motivation in learning at the school (28% learners from Group 1 and 33.9% learners from Group 2).

3.2.1.3. Learners' class ranks

The next table shows the class ranks of the learners at the end of the academic year. The learners were ranked by their homeroom teachers based on their overall academic achievements in all subjects taught in the school. Since there were some learners excluded from the study, the rank cannot be put in individual order (i.e., 1, 2, 3, …, n). Therefore, the ranks were categorized as seen in Table 10 below.

Table 10. Learners' class ranks at the end of academic year

Variables	Categories	Group 1		Group 2	
		n	%	n	%
Class rank	Rank 1-5 (1)	17	20.7	13	23.2
	Rank 6-10 (2)	14	17.1	13	23.2
	Rank 11-15 (3)	13	15.9	15	26.8
	Rank 16-20 (4)	17	20.7	10	17.9
	Rank > 20 (5)	21	25.6	5	8.9

3.2.2. Pre-post analysis

As mentioned in the methodology section, to see the progress of the learners' writing scores, pre-post approach was employed. For the pre and post scores, the average scores of the first three writings (pre) and the average scores of the last three writings (post) were used. Then, to get a better observation of the learners' progress, the average scores of the middle three writings (mid) were also used in the analysis. To test the normality of the distribution of the data, *Kolmogorov-Smirnov* test were performed. Then,

Levene's test was also carried out to test the homogeneity of the data. When the data were normally distributed and homogenous, then ANOVA and *independent t-test* were performed. In contrast, when the data were not normally distributed and not homogenous, the data were analysed using non-parametric tests namely *Mann Whitney* and *Kruskal-Wallis* tests. All tests were carried out using SPSS 22.0.

3.2.2.1. Normality test

This test was carried out to compare the distribution of the data in this study to the standard normal distribution. *Kolmogorov-Smirnov* normality test was carried out using significant value (α) = 0.05. If the α value > 0.05 then, the data is normally distributed. However, the α value < 0.05 means that the data is not normally distributed. The results of the test is shown in Table 11 below.

Table 11. Results of *Kolmogorov-Smirnov* test

Parameters	**Group**	**Test**	**Test Statistic**	***Sig.***	**Distribution**
Times of test (pre, mid, post)	Group 1	-	0,143	0,000	Not normal
	Group 2	-	0,114	0,000	Not normal
Gender	Group 1	Pre	0,263	0,000	Not normal
		Mid	0,117	0,000	Not normal
		Post	0,175	0,000	Not normal
	Group 2	Pre	0,108	0,158	Normal
		Mid	0,115	0,061	Normal
		Post	0,118	0,050	Normal
Group	-	Pre	0,153	0,000	Not normal
	-	Mid	0,149	0,000	Not normal
	-	Post	0,130	0,000	Not normal

The table shows that only the data from the scores of Group 2 has significant value (α) value > 0.05 which means that the data from this category are normally distributed. This means that the data from the other parameters are not normally distributed. This is predictable especially from the Group 1 since most of the students had a low score in the beginning, which made the data right skewed.

3.2.2.2. Homogeneity test

After performing the normality test, homogeneity test was performed on the parameter that has normal distribution using homogeneity of variance test (*Levene's test)* in order to find out whether the data in the parameter have variance that is homogeneous. The test was also carried out using SPSS 22.0 with significance level (α) = 0.05. Data is homogenous when the significance level > α. Conversely, data is not homogenous when significance level < α. The results of the test is shown in Table 12 below.

Table 12. Results of homogeneity tests of Group 2

Group	Parameter	*Levene's Test*	*Sig.*
2	Pre	0.168	0.683
	Mid	0.607	0.439
	Post	0.057	0.813

The table shows that the data from all parameters indicate that significance level > 0.05. This shows that the assumption of normality was satisfied for all parameters in the data from Group 2.

Consequently, from the normality and homogeneity assumption testing it can be decided that *Mann-Whitney U* test will be run to determine if there were differences in the scores based on gender in Group 1. The same test will also be run to determine if there were difference in the scores based on the learners' groups (Group 1 and Group 2). On the other hand, to determine the overall difference between pre, mid, and post in both groups, *Kruskal Wallis* was run. Finally, since only the scores from Group 2 are normally distributed, *independent t-test* will be run to determine if there were any differences between pre, mid, and post of that group.

3.2.3. Regression analysis

A regression analysis was performed for both groups with forced entry including initial writing proficiency, age of acquisition, motivation and gender as predictors to predict the performance on the writing test. Table 13 shows that in case of Group 1, initial writing proficiency and age of acquisition were significant predictors, the latter contributed negatively to the gains. This mean earlier acquisition leading to better gain. Significant regression equation was found ($F(4, 81) = 36.88$, $p = .000$) with an explained variance R^2 of 65%. In Group 2 only initial writing proficiency was found as a significant

predictor. In this analyses too, a significant regression equation was found F(4,55) = 76.77, p .= 000 with an explained variance R^2 of 85%.

Table 13. Multiple regression analyses on the writing scores

	Group 1			Group 2		
	B	SE B	β	B	SE B	β
Initial writing proficiency	.722	.065	.790*	.943	.055	.924*
Age of acquisition	-.034	.015	-.161*	.005	.013	.023
Motivation	.044	.082	.037	-.009	.154	-.009
Gender	.021	.074	.020	.036	.056	.038
R^2			0.65		0.85	
F			36.88		76.77	

3.2.3.1. The effect of class rank

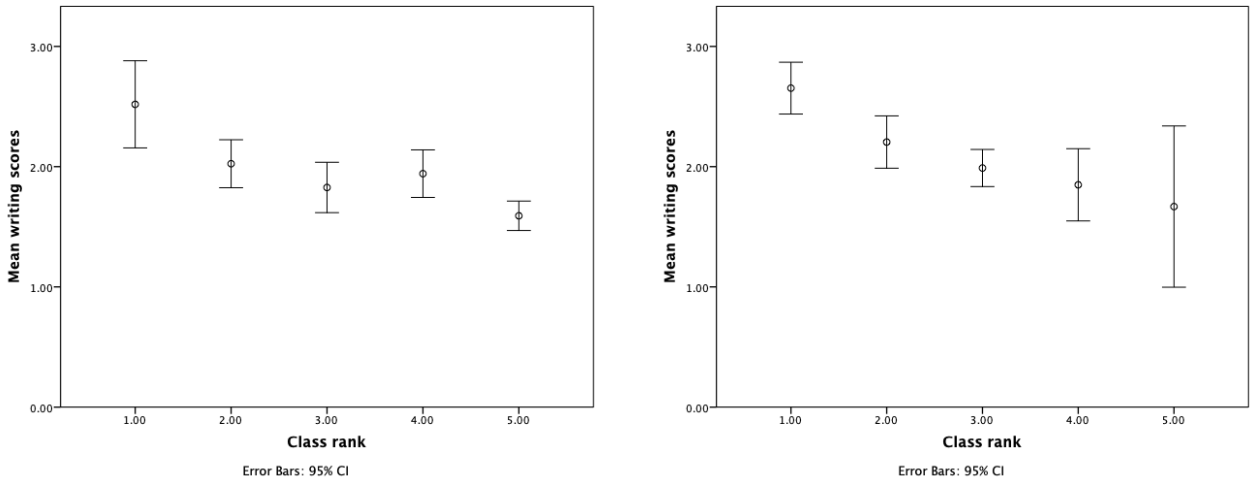
a. Group 1

A one-way between-subjects ANCOVA was calculated to examine the effect of class rank on the writing scores controlling for the initial writing proficiency (covariate). Class rank did show a significant difference in terms of writing scores F (4, 76) = 4.613, p =.002. Initial writing proficiency was significantly related to the gains F (1, 76) = 85.776, p = .000. Post-hoc Bonferroni tests showed that there was a significant difference between class rank 1 and every other rank ($p < .05$) in terms of writing scores, while the rest of the ranks did not differ significantly.

b. Group 2

Initial writing proficiency showed a strong significant relationship with the gains (r= .925, p=.000). The one-way between-subjects ANCOVA showed no significant effect of class rank on the gains when controlled for initial writing proficiency (covariate), but the difference was significant when the covariate was excluded F (4, 51) = 10.649, p = .000. The Bonferroni post-hoc test showed significant difference between class rank 1 and every other level, while the rest of the ranks did not differ significantly (see Figure 8).

Figure 8. Writing scores according to class rank (Group 1 to the left, Group 2 to the right)



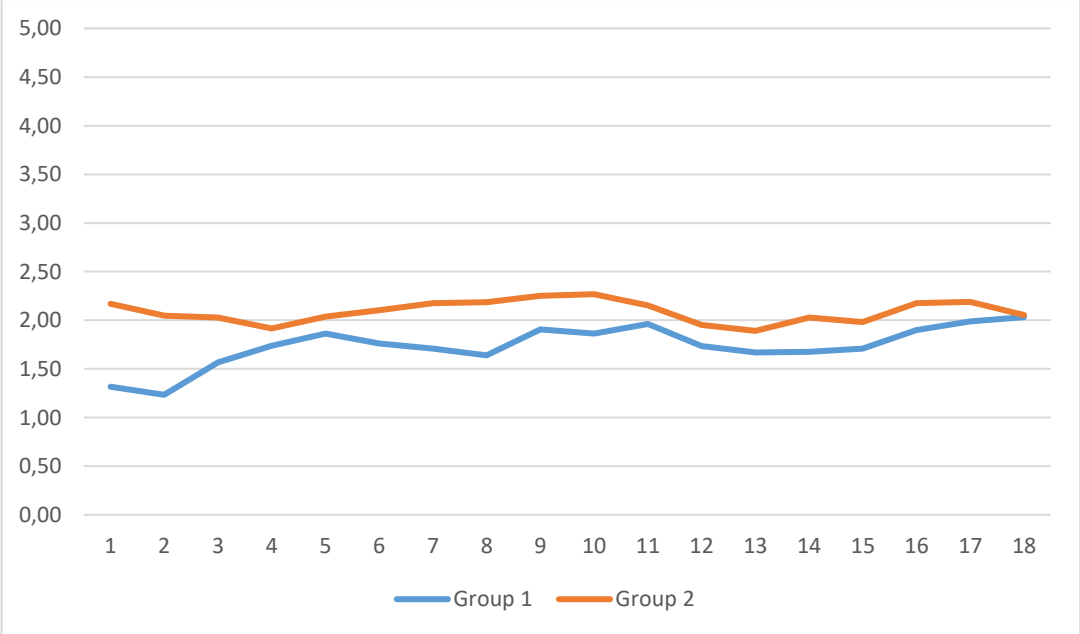
3.3. L2 development

The third study investigates the extent to which the learners' texts change overtime from a Dynamic Usage Based perspective in terms of holistic scores, both at the group level and at the individual level. It also seeks to find out whether the learners show variability over time and whether more variability can be related to more gains.

3.3.1. Descriptive group analysis

Figure 9 below shows how the learners in Group 1 and Group 2 develop overtime during a period of one year. The x axis shows the writing sessions while the y axis shows the score of the learners.

Figure 9. Development of score averages in Group 1 and Group 2



The graph shows that initial score average of the learners from Group 1 is lower than learners from Group 2. However, the scores of the learners in group show an increasing trend with noticeable fluctuation in the first half, while the scores of the learners in Group 2 tend to form a plateau throughout the period. The next chart will demonstrate the development of score averages of male and female learners in Group 1 during 1 year period.

3.3.2. Difference tests between pre, mid, and post of both groups

Kruskall Wallis test was run to determine if there were any difference in the scores of pre, mid, and post in both groups with significance level (α) = 0,05. The results can be seen in Table 14.

Table 14. Results of *Kruskall Wallis* test

Group	Test Statistics	*Sig.*	Significance
1	88,047	0,000	Significant
2	2,198	0,333	Not significant

The results show that there is a significant difference in the scores of Group 1 but not in the scores of Group 2. This can be seen from the significance value of the 1st group which is < 0.05. To see which pairs of tests were significantly different, *a comparison test* was run. Table 15 shows the results of the test.

Table 15. Results of Comparison test

Sample1-Sample 2	Test statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Pre-Mid	-105.665	13.757	-7.681	.000	.000
Pre-Post	-107.264	12.218	-8.779	.000	.000
Mid-Post	-1.599	12.218	-.131	.896	1.000

Each node shows the sample average rank of test.

It can be seen from table 15 above that, in the first group, there is significant difference in the scores between start and mid as well as between the start and end. However, there is no significant difference in the scores between mid and end. The results, therefore, suggested that the first group showed significant improvement in the first semester but not in the second semester. Next, the following table shows the score averages of the learners' writings. As mentioned earlier, the average scores of three writings in the beginning (pre), three writings in the middle (mid), and three writings in

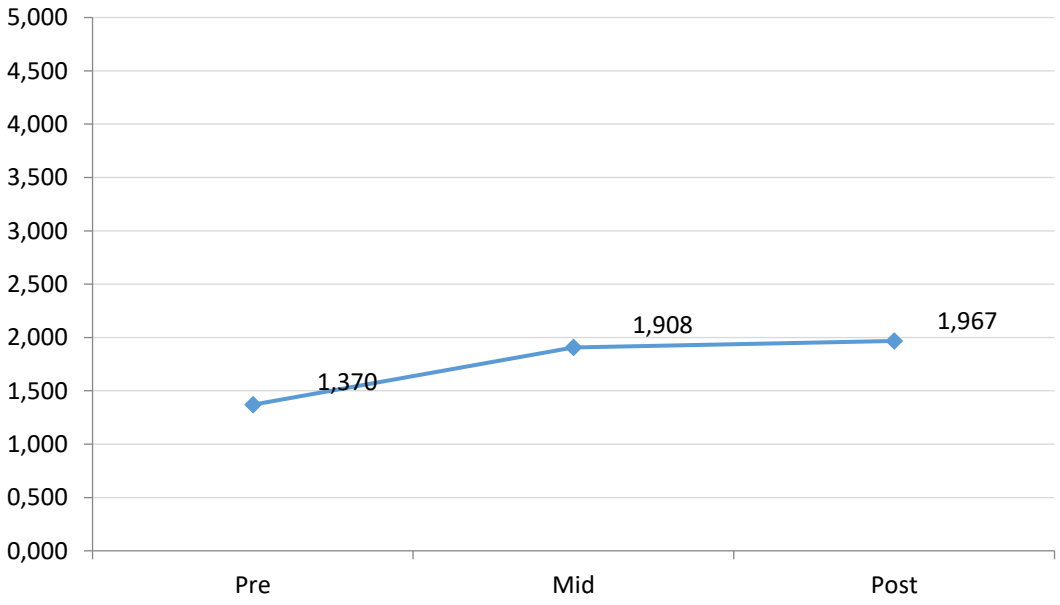
the end (post) were used to determine the significance of the learners' development in the pre-post analysis.

Table 16. Score average pre, mid, and post

Group	Male			Female		
	Pre	**Mid**	**Post**	**Pre**	**Mid**	**Post**
1	1.478632	1.950855	2.014957	1.271318	1.868217	1.924419
2	2.006667	2.093333	2.053333	2.145161	2.327957	2.209677

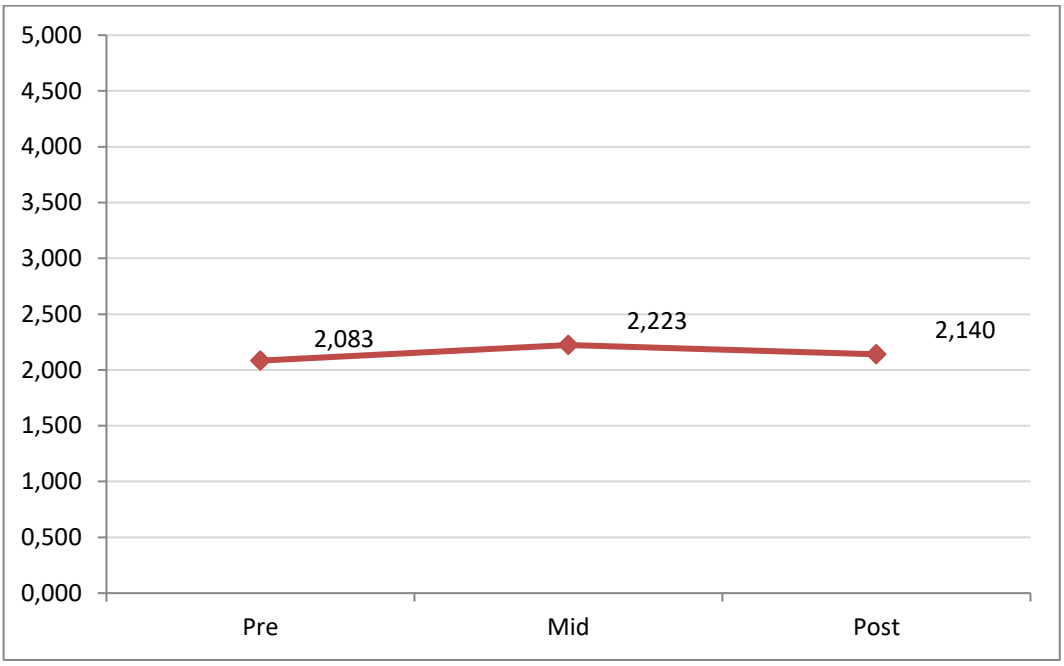
Table 16 shows that, overall, the average scores of learners from Group 2 are higher compared to the scores of learners from Group 2 in pre, mid, and post. It can also be seen from the table that there are some developments in the writing scores during the one-year period. Figure 10 show the development of the writing scores of the learners from Group 1.

Figure 10. Average scores of Group 1



The graph shows that there is a noticeable increase between pre and mid. However, there is only a slight increase between mid and post. In Group 2, however, based on Figure 11 below, it can be seen that there are no significant differences between pre, mid, and post. There is a slight increase between pre and mid. However, there is a slight decrease between mid and post.

Figure 11. Average scores of Group 2



3.3.3. Difference tests between groups

To determine if there were any significant difference in pre, mid, and post scores between the Group 1 and Group 2, *Mann Whitney test* was performed with significance level (α) = 0.05. The results are presented in Table 17 below.

Table 17. Results of difference tests between groups

Test	Test Statistics	*Sig.*	Significance
Pre	-7.705	0.000	Significant
Mid	-3.980	0.000	Significant
Post	-2.754	0.006	Significant

The table shows that the significance level for all tests were < 0.05. Therefore, it can be concluded that, overall, the average scores of Group 2 is significantly higher than Group 1.

3.3.4. Variability in individuals

Two studies (Lowie & Verspoor, 2019; Huang et al., 2020) have found relation between degree of variability, operationalized as the CoV and gains in L2 proficiency level. In the current study, the degree of variability in L2 writing was also operationalized as the coefficient of variation (CoV), which is the standard deviation divided by the mean

performance on the writing test. Another variable was created which is the difference between the performance on the writing test and the initial writing proficiency and denotes how much students gained in the English class. A positive significant correlation was found for Group 1 between CoV and gains: r=-.396, p=.000. Regarding Group 2, the correlation was non-significant and negative between the two measures: Gains r=-.25, p=.063.

A regression analyses was performed with gains as the outcome variable and CoV, class ranking and initial proficiency as predictors. Table 18 shows that CoV and class ranking were significant predictors of performance on the writing test in Group 1. Significant regression equation was found ($F(3, 81) = 280.81$, $p = .000$) with an explained variance R^2 of 91%. In Group 2 the same variables, CoV and class ranking turned out to be significant predictors of students' performance and a significant equation was found ($F(3, 55) = 3.014$, $p = .038$), though the explained variance R^2 was very low, only 0.9%.

Table 18. Regression analysis results

	Group 1			**Group 2**		
	B	**SE B**	**β**	**B**	**SE B**	**β**
Initial writing proficiency	.046	.052	.052	-.309	.155	-.370
Class ranking	-.050	.013	-.145*	-.115	.053	-.371*
CoV	1.25	.05	.937*	-3.49	1.413	-.350*
R^2			0.91		0.09	
F			280.81		3.015	

3.4. Pidginization

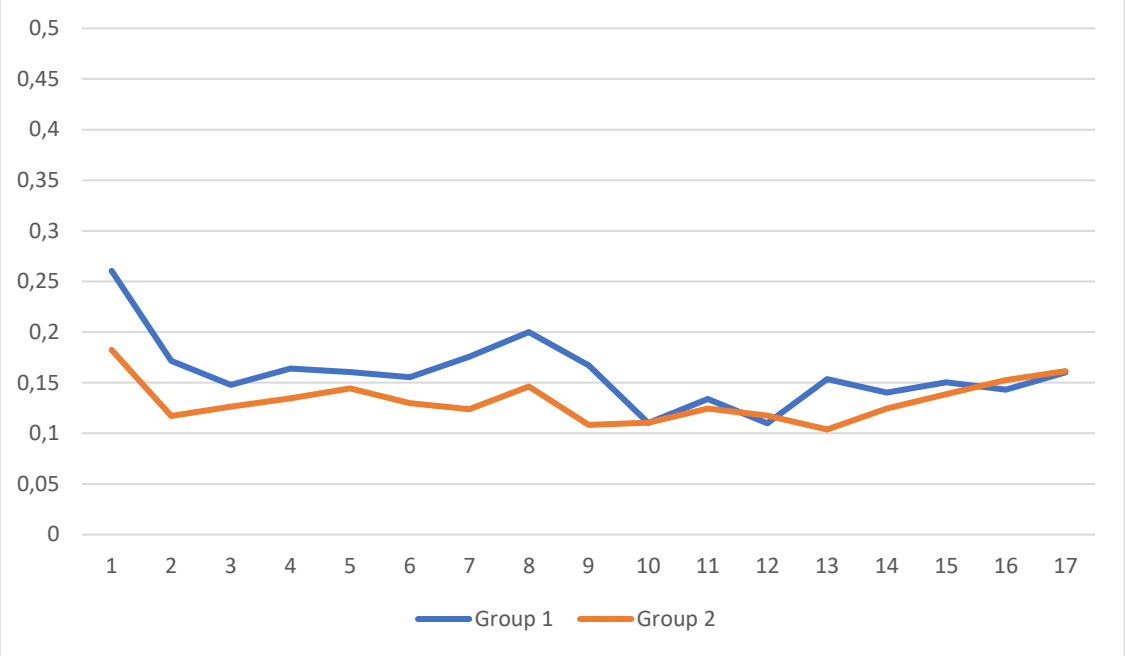
This fourth study is the final study in this dissertation. It seeks to explore the extent of the pidginization features in the learners' L2. Moreover, it also attempts at finding out the distribution of the features in the learners L2.

3.4.1. Development of pidginization features

Figure 12 below shows the development of ratio of pidginization features (P-forms) to the number of words in each session produced by the learners in Group 1 and Group 2 during the observation period. The x axis shows the writing sessions while the y axis shows the ratio. A higher ratio means that the learners produced a greater number of P-forms. The figure shows that the percentage of the pidginized forms in Group 1 decrease rapidly in the first three writing sessions. However, it tends to stagnate in the

following sessions. In Group 2, although there is a slight decrease in the beginning, the line shows a rather steady development.

Figure 12. P-Form ratio of Group 1 and Group 2



To see whether there is a significant development of the ratio of the p-form in both groups, we ran a paired sample t-test. As mentioned in the methods section, the first session in year 1 was not included in the analysis because the texts were very short and contained memorized forms and showed little creative language use. In the first session, the learners only introduced themselves (i.e., name, age, birth date, place of origin, etc.) which requires considerably less linguistic complexity and L2 proficiency. The average of sessions 2-3 and the last two tests were taken as pre and post scores to even out variability and effect of writing topic.

As can be seen in Table 19, Group 1 improved significantly during the observation period with fewer P-forms ratio in the last two sessions ($t(10)=2.496$, $p=.034$). Their ratio in the end (M = .152, SD = .026) is lower than their ratio in the beginning (M = .216, SD = .071).

Table 19. Paired sample t-test results of Group 1

Variable	pre		post		t(10)	*p*	Cohen's D
	M	SD	M	SD			
P-form ratio	.216	.071	.152	.026	2.496	.034*	1.313

*$p<.05$

Group 2, on the other hand, did not show any significant improvement during the observation period (t(10)=2.496, p=.034). There is no significant difference in terms of their ratio in the end (M = .152, SD = .026) to their ratio in the beginning (M = .216, SD = .071) as seen in Table 20 below.

Table 20. Paired sample t-test results of Group 2

Variable	pre		post		t(10)	*p*	Cohen's D
	M	SD	M	SD			
P-forms ratio	.150	.033	.157	.094	-.659	2.262	-.115

Then, to see whether there is any difference between Group 1 and Group 2 in terms of the P-forms ratio, an independent-samples t-test was run. The results in Table 21 show that, in the beginning, the learners in Group 2 (M = .150, SD = .034) produced significantly fewer P-form ratio (t(10) = 2.653, *p* = .019) when compared to the learners in Group 1 (M = .216, SD = .071).

Table 21. Independent t-test results of pre scores between Group 1 and 2

Variable	Group 1		Group 2		t(10)	*p*	Cohen's D
	M	SD	M	SD			
P-form ratio	.216	.071	.150	.034	2.653	.019	1.261

*$p<.05$

However, in the end of the observation period, there was no significant difference in terms of the P-form ratio (t(10) = -.463, p = .668) between learners in Group 1 (M = .151, SD = .027) and Group 2 (M = .157, SD = .030) as seen in Table 22 below.

Table 22. Independent t-test results of post scores between Group 1 and 2

Variable	Group 1		Group 2		t(10)	*p*	Cohen's D
	M	SD	M	SD			
P-forms ratio	.151	.027	.157	.030	-.436	.668	-.195

The next independent t-test analysis was run to compare the P-form ratio of Group 1 in the end of the observation period and the P-form ratio of Group 2 in the beginning of the observation period.

Table 23. Independent t-test results of Group 1's post score and Group 2's pre scores

Variable	Group 1		Group 2		t(10)	*p*	Cohen's D
	M	SD	M	SD			
P-forms ratio	.151	.027	.150	.034	0.137	.892	.062

Table 23 above shows that there is no significant difference between the P forms ratio (t(10) = .137, p = .892) in the end of the post scores of Group 1 (M = .151, SD = .027) and the pre scores of Group 2 (M = .150, SD = .034).

The analyses show that although the learners started differently, with the learners in Group 1 having significantly more P-forms than in Group 2, the learners in Group 1 equalled out at the end of the observation period. This was indicated in the independent t-test results between the post scores of P forms ratio of Group 1 and the pre scores of p forms ratio of Group 2. Moreover, the post scores of P forms ratio from both groups show that they are similar.

3.4.2. Types of pidginization features

To find out the distribution of pidginization features in the learners L2, we counted each feature in the writing samples. The codes are as follows: basic pidgin negation (1), lack of inversion (2), lack of auxiliary (3), lack of possesive inflection (4), lack of inflectional morphology (5), L1 mix (6), L1 forms and constructions (7).

Table 24 below show that L1 forms and constructions (56%) made up the majority of the features found in the writing samples from Group 1, followed by a lack of inflectional morphology (20%) and lack auxiliary (13%). Instances of L1 mix (7%) were also found but mostly in the earlier sessions. Basic pidgin negation (2%) and lack of possessive inflection (2%) were scarcely found while we only found two instances of lack of inversion in questions.

Table 24. Occurrences of pidginization feature types in Group 1 and 2

No.	Type	Group 1	%	Group 2	%
1	basic pidgin negation	27	2.2	33	2.8
2	lack of inversion	1	0.1	0	0
3	lack auxiliary	166	13.4	134	11.4
4	lack possessive inflection	23	1.8	4	0.3
5	lack inflectional morphology	247	20	248	21.1
6	L1 mix	87	7	68	5.8
7	L1 forms and constructions	686	55.4	689	58.6
	Total	1237		1176	

The figure is superficially similar in group two. L1 forms and constructions (59%) was also the majority of the features produced by the learners in Group 2, followed by lack of inflectional morphology (21%) and lack auxiliary (11%). Instances of L1 mix (6%) were also found. Basic pidgin negation made up only 3% of the features found. Only 4 instances of lack of possessive inflection were found while no instances of lack of inversion in questions was found. Table 25 below shows some examples of the different types of pidginization.

Table 25. Examples of pidginization features

No.	Type	Group 1	Group 2
1	basic pidgin negation	… the pupils very kindness, excaiting and *no angry* anymore	… so if *no play foot ball,* I play game.
2	lack of inversion	… so I don't know *what must I do.*	-
3	lack auxiliary	In over there, I feeling happy attogh *the place hot*.	… and I like milk because *milk good for body.*
4	lack possessive inflection	We are giving name to *we club,* (name).	… *they movie* very funny.
5	lack inflectional morphology	Last holiday, *I go* to (name), (name) and living plaza.	If swimming, *I have study* from 2nd grade.
6	L1 mix	Tomorrow morning, I am *terlambat.*	… me too like math and *fisika* lesson.
7	L1 forms and constructions	I *must can drawing*.	… with my friend *I can play long.*

3.5. Summary

In the first study on interaction, the examination on the learners' interactions shows that the learners in the Group 1 produced noticeably fewer turn takings (n=107) than the learners in Group 2 (n=286). Although the number of turn takings differ quite significantly, the difference in terms of the percentage of turn takings containing non target like (NTL) utterances is not significant. The turn takings in Group 1 contains 54% (n=54) NTL utterances while the turn takings in Group 2 contains 43% (n=124) NTL utterances. In terms of the interactional features, the results from the analyses indicated that the learners produce a very small number of interactional features that are reportedly important for language learning. None of the NTL utterances produced by the learners in Group 1 resulted in feedback and the only 3 instances of modified output were self-initiated. Moreover, 2 out of these 3 modified outputs were still NTL. In the second group, only 7 negative feedbacks, of which 2 were NTL, were produced by the learners as a response to triggers produced by their conversation partners. There were also 6 self-initiated modified outputs in the second group, one of which was NTL.

In the second study, on predictors for L2 development, a regression analysis was performed for both groups with forced entry including initial writing proficiency, age of acquisition, motivation and gender as predictors to predict the performance on the writing test. Results show that in the case of Group 1, initial writing proficiency and age of acquisition were significant predictors, the latter contributed negatively to the gains. This means earlier acquisition leading to more gains. In Group 2, on the other hand, only the initial writing proficiency was found as a significant predictor. Scholastic aptitude in terms of class rank did show a significant difference in terms of writing scores in Group 1. However, in Group 2, one-way between-subjects ANCOVA showed no significant effect of scholastic aptitude in terms of class rank on the gains when controlled for initial writing proficiency (covariate), but the difference was significant when the covariate was excluded.

In the third study, on L2 development over time, the result shows there is a significant difference in the scores between start and mid as well as between the start and end in the first group. However, there is no significant difference in the scores between mid and end. The results, therefore, suggested that the first group showed significant improvement in the first semester but not in the second semester. In contrast, in Group 2 there was no significant difference between pre, mid and post scores. This means that the learners in Group 2 did not make any significant progress during the one-year period

despite the fact that their overall average score is higher than the first group. Then, a regression analyses was performed with gains as the outcome variable and CoV, class ranking and initial proficiency as predictors. Results show that the CoV was a significant predictor of performance on the writing test in both Group 1 and Group 2.

Finally, the results in the study on pidginization features shows indications of pidginization in the learners L2. In the paired-samples t-test, learners in Group 1 show that they improved significantly by producing fewer pidginization features overtime. Several runs of independent t-tests show that although the learners started differently, with the learners in Group 2 having a significantly better ratio than the learners in Group 1, the learners in Group 1 equalled out in the end of the observation period. This was indicated in the independent t-test results between the post scores of P forms ratio of Group 1 and the pre scores of P forms ratio of Group 2. Moreover, the post scores of P forms ratio from both groups show that they are similar. We also counted each type of pidginization feature and found that both groups produced a rather similar percentage of the features. L1 forms and constructions made up the majority of the features, followed by a lack of inflectional morphology and lack of auxiliary. Instances of L1 mix were mostly found in the earlier sessions. Basic pidgin negation, lack of possessive inflection, and lack of inversion in questions were scarcely found in the groups.

In the next chapter we will discuss the findings and relate them back to the literature.

CHAPTER 4
DISCUSSION

This thesis set out to trace the L2 English development in two cohorts of 82 and 56 students at a Pesantren in Indonesia over the course of one academic year. The type of instruction for English, explained in detail in Chapters 1 and 2, consisted mainly of a lexical approach in that word lists were given and discussed, and the young learners were asked to use these words in English in peer interaction in their English weeks. All in all, the instructional approach consisted of about 90% peer interaction and a few hours of regular English classes during the school week. However, there was little to no authentic input and the words that were taught were taught thematically but in isolation. The set up reminds us vaguely of the Mester et al.'s (1989) experiment which was based on the previously rejected proposal of Bickerton and Givón, to see how speakers of different first languages would develop a language based on L2 words and interaction. The aim of this dissertation was also to find out how the learners' language would develop over time. Taking a dynamic usage based view of language learning, we assumed that with so little authentic input and so much repetition of learners' non-target utterances the learners might create their own version of English, which would eventually stabilize and could be considered a Pidginized version. Four interrelated studies were devised to test this hypothesis. This chapter presents the results of these four studies: learners' interaction, individual differences, L2 development and Pidginization.

4.1. Learners' interaction

Study 1 concerned peer interaction, particularly in terms of the interactional features which reportedly promote L2 acquisition. We examined samples of the learners' interaction for the interactional features: corrective feedback in the forms of recasts, clarification requests, explicit corrections and modified output (as response to feedback or self-initiated), including triggers (i.e., errors produced by the learners during interaction). Moreover, this study looked at how learners from the different years differ in terms of the aforementioned interactional features.

The examination shows that the learners in the Group 1 produced noticeably fewer turn takings (n=107) than the learners in Group 2 (n=286). Although the number of turn takings differ quite significantly, the difference in terms of the percentage of turn takings containing non target like (NTL) utterances is not much different. The turn takings in

Group 1 contain 54% (n=54) NTL utterances while the turn takings in Group 2 contain 43% (n=124) NTL utterances. In terms of interactional features, the results from the analyses indicated that the learners produce a very small number of interactional features that are reportedly important for language learning. None of the NTL utterances produced by the learners in Group 1 resulted in feedback and the only 3 instances of modified output were self-initiated. Moreover, 2 out of these 3 modified outputs were still NTL. Here is an example of the self-initiated modified output taken from Group 1 (underlined). It can be seen that the learner tried to correct herself. However, the modified output is somehow still inaccurate in the English standard.

F1A: *[Every years in this boarding, of course, there are…there were…there were a big agenda for us.]*

In the second group, only 7 negative feedbacks, of which 2 were NTL, were produced by the learners as a response to triggers produced by their conversation partners. There were also 6 self-initiated modified output in the second group, one of which was NTL. The following excerpt from the transcript taken from the second group shows how poor the feedback and the modified output were (underlined). It can be seen how one learner was attempting to correct his partner, which results in rather confusing exchanges.

M2B: *[Old? Uh… I was twelve years… twelve years old.]*
M2A: *[Ah, twelve years old.]*
M2B: *[Yes.]*
M2A: *[The youngest? Youngest?]*
M2B: *[No, just not youngest but younger.]*
M2A: *[Ah yes, the younger.]*
M2B: *[Younger than you and the youngest is…]*
M2A: *[Yes. I am oldest. I am older. You are youngest.]*
M2B: *[Yes. Yes.]*

These findings indicate that the peer interaction among the learners in the pesantren have few to no interactional features that can promote language learning. Although there were a few examples of feedback which resulted in modified output in Group 2, the quality is still questionable.

This outcome is contrary to previous studies by Mackey et al. (2003) and Shehadeh (1999), which suggested that peer interaction encourages more output-promoting feedback and more accuracy in the utterances. However, age may also play a

role in the case of pesantren because as Oliver's (1998) study suggested, child learners tend to produce fewer interactional features necessary for L2 learning compared to adult learners. In a more recent study, Oliver et al. (2017) compared two groups of young learners (5-8 years and 9-12 years) and found that, in some topics of the task, the older group of learners tend to produce less negotiation of meaning because they 'simply wanted to get the task done' (2017: 8). This may also provide an explanation for the findings of the current study since the age range of both Group 1 and Group 2 is about the same as the second group in the study by Oliver et al. (2017).

The only encouraging signs found in this study is in the difference between the groups in terms of turn taking, in which Group 2 almost tripled the number of turn takings produced by Group 1. This may indicate that they have become more fluent as they progressed, which confirms the statement arguing that interaction can improve fluency and automaticity (Lyster & Sato, 2013; DeKeyser, 2017a). The study by Xu et al (2019) may also provide a possible explanation for the lack of the interactional features. They found that learners were hesitant in providing corrective feedback to their peers. In their study, the learners preferred using recasts instead of prompts and explicit correction in their infrequent corrective feedbacks. This is also shown in the findings of the current study, especially in Group 2. However, in the current study, the number is still considerably low. Another possible explanation for this is that the learners in the pesantren have relatively poor English proficiency to begin with and have not been exposed to target-like language so they might not have an adequate English repertoire to notice non-target like utterances and, consequently, to provide feedback.

4.2. Individual differences

Study 2 examined the effect of individual differences such as gender, language background, motivation and scholastic aptitude on the learners' L2 writing development. It also examined possible differences between the first group and the second group. From a DUB perspective, initial conditions of the learners are very important and learners are expected to have different individual trajectories in their development. Learners' personal and linguistic background such as L1, scholastic aptitude, motivation, etc. are assumed to serve as predictor variables, which interact in complex manners and determine the acquisition of the L2.

The students filled out an extensive Language History Questionnaire and the results showed that almost all learners were multilingual. Most had Sundanese as their L1

and then learned the lingua franca, Indonesian, early on. Most of the learners in Group 1 and 2 rated their self-report language learning ability 4 and 5 in the scale of 1-7. Indonesian is the language the majority of the students are most comfortable with in terms of writing, while Sundanese is the language, they are most comfortable with in terms of speaking. The results were so complex that it was not really possible to categorize the learners in a few defined groups so not all information from the results were considered in further analyses.

For motivation, the learners were asked to write a reflection in about 100-200 words in their L1, on their motivation to enroll in the school. Then, the learners' reflections on their motivation were scored based on Self-Determination Theory (SDT) (Utvær & Haugan, 2016) from which a scale was developed by Deci and Ryan (2008), and Ryan and Deci (2009). The result shows that most of the learners have autonomous motivation and show indications of internalization. Only a small number of learners had low motivation in learning at the school.

To operationalize scholastic aptitude, we took the learners' academic report showing class rank. The documents of the learners' academic report were provided by the school with the permission of the principal and the academic counsellor. They are categorized based on the following order; learners who belong to the top 20% in their class were coded 1, the next 20% in their class were coded 2, and so on. Since every learner has his/her own rank, the number of participants in the scholastic categories was almost equally distributed.

For L2 writing development, this study uses free writings in the form of narratives. The writing sessions were carried out over a one-academic-year period. The learners participated in 18 writing sessions in total, conducted once every other week in their English classes. Every writing session lasted approximately 20 minutes. To assess the general proficiency of the learners, their writings were holistically scored. A group of five raters developed the scoring criteria and agreed on six levels of proficiency (0–5). From the discussion, the raters also agreed on half scores (e.g., 1.5, 2.5, etc.) for texts that have features of two different scores.

A regression analysis was performed for both groups with forced entry including initial writing proficiency (operationalized by the average of the first three scores), age of acquisition, motivation and gender as predictors to predict the performance on the writing gains. It was revealed that in the case of Group 1, initial writing proficiency and age of acquisition were significant predictors. Age of acquisition contributed negatively

to the gains which means the earlier they started learning English, the better their gains were. In Group 2, on the other hand, only the initial writing proficiency was found as a significant predictor. The fact that initial proficiency is a strong predictor in this study is in line with Verspoor et al. (2015), which even found that initial proficiency is the strongest contributor to gains in their study.

Gender and motivation on the other hand are not strong predictors in the groups. This is in contrast to some studies on gender (e.g., Oxford, 1993; Young & Oxford 1997) and motivation (Oxford, Nyikos & Ehrman, 1988; Gardner, 1985; Saville-Troike & Barto, 2016), which have indicated they can play a significant role on how language learners develop. Studies in the role of gender in L2 acquisition have generally suggested that females are better language learners because they tend to have a more positive attitude towards L2, show better integrative motivation, and utilize a wider range of learning strategies (Oxford, Nyikos & Ehrman, 1988). Motivation has also been widely considered as one of the most prominent factors affecting L2 acquisition. One of the leading researchers in this area, Gardner (1985), found that higher motivation could result in more desire and effort to achieve learning goals. It also leads to a more positive attitude in the learning process. Saville-Troike and Barto (2016) even claim that motivation is the second most significant predictor after aptitude in the success of second language learning. However, this claim may not always be correct for all ages or stages of development. For instance, in a study conducted by Verspoor, de Bot, and Xu (2015), motivation had a significant contribution in L2 development in the first group in year 1 but not the other.

Another important variable that was analysed in the regression analysis was scholastic aptitude in the form of class rank. It has been long reported to be an important factor in the success of L2 learning (e.g., Carroll, 1981, 1990; Skehan, 1989). However, only few have been conducted in an interactionist perspective (e.g., Mackey et al., 2002; Trofimovich et al., 2007; Goo, 2012; Révész, 2012). For example, Trofimovich et al. (2007) found that learners' working memory, phonological memory, analytical ability, and attention control are highly correlated to the learners' ability to notice and benefit from recasts. As mentioned in the methods section, scholastic aptitude in this study was measured in terms of class rank. Interestingly, scholastic aptitude did show a significant difference in terms of gains in Group 1. However, in Group 2, one-way between-subjects ANCOVA showed no significant effect of scholastic aptitude in terms of class rank on the gains when controlled for initial writing proficiency (covariate), but the difference

was significant when the covariate was excluded. This is in line with Verspoor et al. (2015) who also used scholastic aptitude as one of the predictors in their study. Similar to motivation which has been mentioned earlier, they found that while it is a strong predictor in the first group, it does not serve as a significant predictor in the second group. In addition, it is interesting to note that the participants in their study are approximately the same age as the ones in the current study. In an interactional context, a few studies have suggested that a higher aptitude can be beneficial for interaction which eventually can lead to a better L2 acquisition (e.g., Li, 2013). The fact that in Group 1 scholastic aptitude plays a role and in Group 2 does not may also suggest that initially these differences contribute to faster learning, but as the learners' development stabilizes, nothing affects their development much anymore. One reason may be that once people have enough to communicate, they do not improve anymore (Schumann, 1978).

4.3. L2 development

Study 3 attempted to explore the extent of the development of English learners at the pesantren and whether the learners show variability overtime. Moreover, it also tried to find out whether there was any difference between Group 1 and Group 2 as Group 1 were beginners and Group 2 already had one year at the pesantren. Before discussing the findings of this study, it is important to understand that L2 development can be regarded as a dynamic process of change (Larsen-Freeman, 1997; De Bot, 2008; and Verspoor, De Bot and Lowie, 2011). The dynamics of such process cause changes to be non-linear with a significant extent of variability (within systems) and variation (among systems). As De Bot and Larsen-Freeman (2011) put it, sometimes a system "changes continuously, sometimes discontinuously, even chaotically". Systems, however, have a tendency to move towards preferred states, called attractors. Therefore, language development cannot be represented by a straight linear continuum.

To explore actual development, it is therefore not enough to do simple pre- and post-tests but to trace learners over time. In the current study, each learner in the analysis produced about 18 texts in English over the course of one academic year. Each text was retyped and scored holistically on the level of development according to the rubric presented in Table 2.5 in Chapter 2 of this dissertation.

To see if there was actual development, operationalized as an increase in scores, we compared group pre-scores with mid scores and final scores, both in Group 1 and Group 2 and as Figure 3.2, 3.3, and 3.4 shows and the statistical analyses showed Group

1 improved significantly in the first half year and then stabilized. Group 2 was significantly better than Group 1 only in the first scores, but after that there were no significant differences. The results, therefore, suggest that the first group showed significant improvement in the first semester but not in the second semester. In contrast, in Group 2 there was no significant difference between pre, mid and post scores. This means that the learners in Group 2 did not make any progress during the one-year period. Although the overall average score of Group 2 is higher than Group 1, this is because Group 2 started off with better scores in the first 6 months than Group 1.

On the whole there was very little progress in Group 2. Simple word counts even show that Group 2 produced slightly fewer words than Group 1. Moreover, their writing did not become more sophisticated in terms of complexity or accuracy nor non-target forms, especially when compared to the later writings from Group 1. Still some learners improved more than others and we checked if those who did also were more variable in terms of ups and downs in their scores.

A regression analyses was performed with gains as the outcome variable and variability measured through CoV, class ranking and initial proficiency as predictors. Results show that the CoV was a significant predictor of performance on the writing test in both Group 1 and Group 2. This finding is in line with the study by Huang et al. (2021) which shows that besides the traditional individual differences, variability has a strong correlation with L2 proficiency gains. In their study, Huang et al. (2021) did a multiple linear regression analyses and found that variability was a strong predictor of gains and final L2 writing proficiency when the initial proficiency of the participants was controlled for. Variability is needed to improve. Several studies also have reported that the degree and pattern of variability can provide an insight into the development of L2 learners (de Bot, Lowie, & Verspoor, 2007; Lowie & Verspoor, 2019; Spoelman & Verspoor, 2010; Verspoor & de Bot, 2021).

The developmental findings were very surprising. Despite numerous hours of peer interaction and an ever-increasing list of words and also regular English lessons, on the whole the learners' English proficiency did not improve much after the first six months and actually stabilized. In several longitudinal studies we have seen that there is a strong increase in proficiency early on, in the first six months and then the curve usually stabilizes (see Verspoor et al., 2015; Rousse-Malpat et al., 2019), but usually there is still progress after that albeit at a slower rate.

These findings support the most important tenet of all usage-based approaches: language is learned from the input and output that the language learner experiences. It was clear that the learners in the pesantren get very little exposure to the target language forms and structures and they mostly get their input from their peers. Moreover, the NTL output they produce rarely gets corrected. These factors may cause the learners to stagnate. This is consistent with the DUB approach which emphasized that changes in learner language which are a consequence of, 1) the frequency of use of L2 in social interaction, and 2) the interaction of constructions in the network in the learner's mind (Verspoor & Behrens, 2011; Roehr-Brackin, 2015). There are no innate systems, so the language learner can only discover and acquire the language through exposure and experience. Frequency of exposure is supposedly the main driver of development. Whatever is heard or used the most will become automated and entrenched patterns in the learner's language. Moreover, it is also possible that the learners lack the urgency to improve since the language that they produce is enough to fulfil their daily communicative needs.

4.4. Pidginization

Study 4 seeks to explore the extent of the fossilization/pidginization features in the learners' L2 in the context of pesantren. Moreover, it attempts at finding out the distribution of the features in the learners L2. This study was particularly inspired by an attempt in 1979 by Derek Bickerton and Talmy Givón, who proposed an experiment in which people speaking mutually unintelligible languages are taught approximately 200 words of English and then placed on an uninhabited island for a year where they would communicate using only the English lexicon while performing agricultural activities (as cited in Master, Schumann, and Sokolik, 1989). Although the proposal was eventually rejected due to potential dangers to the subjects, the concept was later replicated by Mester et al. (1989) in a laboratory setting. This condition is similar to the context of the pesantren in Indonesia, where the type of instruction for English, as described in the earlier chapters, consisted mainly of a lexical approach in that word lists were given and discussed, and the young learners were asked to use these words in their daily communication.

Before discussing the results of this study, it is important to understand the parallel between a DUB perspective and pidginization. A DUB theoretical perspective may inform us about language development that changes over time or stagnates, either with

target norms or non-target forms, the latter being called fossilization in SLA studies. From a DUB point of view, any system may move to an attractor state, where it is likely to remain, which can be related to studies of pidgin and creole languages. Although pidgin has been largely described as a contact language which develops when groups of people who speak different languages attempt to communicate with one another (Richards & Schmidt, 2010), others (e.g., Crystal, 2010) also use the terms "makeshift", "marginal" language, or "mixed languages" to define a pidgin, which to some extent also characterizes learners' language. Moreover, the characteristics of pidgins, i.e., simplifications and transfers of the source languages, are also typical of the language of a learner. Nemser (1974) described these features of learner talk as approximative system while Selinker (1972) used the term interlanguage (IL). Both forms are characterized by a limited system of auxiliary verbs, simplified question and negative forms, and reduced rules for tense, number, and other grammatical categories.

Pidgins have also commonly been observed to have limited lexis, morphology, syntax and a narrow range of use, which can expand and develop when they are used over an extensive period and when their purposes expand. However, they are not 'bad' versions of the source languages but rather highly regularized varieties (see Todd, 1974; Mühlhäusler, 1986; & Romaine, 1988). These features are also commonly produced by language learners and are typically temporary in the language learning process. L2 learners will usually move towards more target like forms in their L2 production as they progress. For example, in a cross-sectional usage-based study on Dutch learners of English, Vries and Verspoor (2010) found that learners' L1 transfer errors go rapidly down between the two beginner levels, level 1 and 2. The decrease in L1 transfer errors is common in the language learning process. However, in some cases, some NTL forms remain and fossilize.

Richards (1974) argued that a fossilized form of the target language could be considered pidginized when learners do not advance beyond this stage. He closely compared pidgin languages and second language acquisition (SLA) and argued that both codes can be described "as an interlanguage arising as a medium of communication between speakers of different languages, characterized by grammatical structure and lexical content originating in differing sources, by unintelligibility to speakers of the source languages and by stability" (Richards, 1974: 77). Some research (e.g., Cancino, Rosansky & Schumann, 1974; Schumann, 1978; and Andersen, 1981) indicate further the similarities between the structures of pidgin languages and interlanguage (see Selinker,

1972; 1992) of L2 learners. Although one may argue that a pidgin arises from language contact between groups of speakers with different languages who are forced to communicate, it is also possible that learners who have the same L1 produce similar features of a pidgin, like in the case of the pesantren in the current study where the learners are forced to abandon their L1 and use the L2 in their daily communication.

In this study, sample texts were examined for the characteristics of pidginization as suggested by Andersen (1981, as cited in McLaughlin, 1987), as well as Schumann (as cited in Johnson & Johnson, 1999). The findings in the study on pidginization features shows indications of pidginization in the learners L2. In the paired-samples t-test, learners in Group 1 show that they improved significantly by producing a lower pidginization ratio overtime. However, when tracked longitudinally (Figure 3.5), the substantial improvement mostly occurred in the first few sessions and then they seemed to stabilize afterwards. When compared to the earlier stages of pidgin languages, these findings may suggest that the pidginization process occurs rapidly early on.

Several runs of independent t-tests show that although the learners started differently, with the learners in Group 2 having a significantly better ratio than the learners in Group 1, the learners in Group 1 equalled out by the end of the observation period. This was indicated in the independent t-test results between the post scores of P forms ratio of Group 1 and the pre scores of P forms ratio of Group 2. Moreover, the post scores of P forms ratio from both groups show that they are similar.

We also counted types of pidginization features and found that both groups produced a rather similar percentage of the features. L1 forms and constructions made up the majority of the features, followed by a lack of inflectional morphology and a lack of auxiliary. Instances of L1 mix were found mostly in the earlier sessions. Basic pidgin negation, lack of possessive inflection, and lack of inversion in questions were also found, but not as often as the other features. However, it is possible that this is because such forms are not commonly produced in great numbers in writing.

The preponderance of pidginization features in the learners' L2 in the pesantren context is in line with the comparative study of pidgin languages and the Alberto's language Schumann (1978), in which he concluded that Alberto's speech was in fact a pidginized version of English. Therefore, it is possible to conclude that the conditions in the pesantren allow for the process of pidginization.

CHAPTER 5
CONCLUSION

The pesantren where this study was conducted was chosen because, as elaborated in the previous chapters, it utilized peer interaction as one of the primary sources of L2 learning. The students are required to speak English one week and Arabic the next week in their daily communication in the hope that it allows students to have extensive practice in the two languages. Such practice is very common in pesantren institutions across the country including Java (e.g., Hidayat, 2007; Aziez, 2016; Al-Baekani & Pahlevi, 2018), Sumatra (e.g., Ritonga, Ananda, Lanin & Hasan, 2019), Sulawesi (e.g., Bin Tahir 2016; Bin Tahir et al., 2017), and even in Papua (e.g., Wekke, 2015). One point that has been consistently reported is the emphasis on peer-interaction and a lexical approach in the language learning practice in pesantren institutions. Observations in the pesantren institutions reveal that learners interact mostly with their peers and very little with more proficient speakers (e.g., teachers). This study aimed to investigate how this peer-to-peer interaction affects the learners' L2 development over time in one academic year in two cohorts, a first-year group with 82 learners (Group 1) and a second-year group with 56 learners (Group 2). This cross-sectional longitudinal design was meant to simulate a two-year developmental path.

Taking a dynamic usage-based (DUB) perspective of language learning, which holds that frequency of exposure and use is the strongest predictor in L2 development, we assumed that with so little authentic input and so much repetition of learners' non-target utterances that the learners might create their own version of English, which would eventually stabilize and be considered a pidginized version. This assumption proved true.

Four interrelated studies were devised to explore pesantren learners' practices and language development. The first study examined the learners' peer interaction, particularly in terms of interactional features, which according to the literature promote L2 acquisition such as turn taking, trigger, negative feedback and modified output. The second study examined the effect of individual differences such as gender, language background, motivation and scholastic aptitude on the learners' L2 writing development. The third study explored development of English learners over time with bi-weekly writing. In the fourth study, the aim was to explore the extent of the fossilization or pidginization the learners' L2 in the context of pesantren. All four studies looked Group 1 and 2 separately and compared them.

In Study 1, samples of learners' interaction were examined for the interactional features which have been reported to promote L2 acquisition (i.e., corrective feedback in the forms of recast, clarification request, and explicit correction; and modified outputs whether they are as response to feedback or self-initiated), including triggers (i.e., errors produced by the learners during interaction). The findings clearly indicate that peer interaction among the learners in the pesantren lacks the interactional features that can promote language learning. Although there were some examples of feedback that resulted in modified output in Group 2, the quality is still questionable.

In Study 2, information on the individual differences, including gender, language background, motivation and scholastic aptitude were gathered through different instruments. A Language History Questionnaire, the learners' reflection on motivation, and academic reports were used for this purpose. For L2 writing development, this study used average scores of the first and last three writings collected in a one-academic-year period. Gains were operationalized as the difference between beginning and end scores. A regression analysis was performed for both groups with forced entry including initial writing proficiency (operationalized by the first score), age of acquisition, motivation and gender as predictors to predict the performance on the writing gains. In Group 1, initial writing proficiency and age of acquisition were significant predictors. Age of acquisition contributed negatively to the gains which means the earlier they started learning English, the higher their gains. In Group 2 only the initial writing proficiency was found as a significant positive predictor. Gender and motivation, on the other hand, were not found to be strong predictors in either group. Scholastic aptitude did show a significant effect on gains in Group 1, but not in Group 2 when initial writing proficiency (covariate) was controlled for. However, scholastic aptitude was significant when the covariate was excluded.

In Study 3, to see if there was actual development, we compared the groups' pre-scores with mid scores and final scores. The statistical analyses showed that Group 1 improved significantly in the first half year and then stabilized. Group 2 was significantly better than Group 1 only in the first scores. The results, therefore, suggest that only in the first semester after enrolling in the pesantren, progress is made. The first group showed significant improvement in the first semester but not in the second semester. In Group 2 there was no significant difference between pre, mid and post scores. This means that the learners in Group 2 did not make any significant progress during the one-year period.

A further regression analyses was performed with gains as the outcome variable and variability measured through CoV, class ranking and initial proficiency as predictors. Results show that the CoV was a significant predictor of performance on the writing test in both Group 1 and Group 2. However, as the CoV does not take time onto account, it may not be the best measure for variability over time and thus this finding needs to be treated with utmost caution.

In Study 4, sample texts were examined for the characteristics of pidginization. The findings show strong indications of pidginization in the learners L2 starting after the first semester in the first year. Learners in Group 1 show that at the beginning they have many more Pidginization forms (P-forms), which are just learner errors than they do later on as they changed significantly by producing relatively fewer P-forms overtime. However, the longitudinal analysis shows that the change occurred mostly in the first few sessions. Several runs of independent t-tests show that Group 1 learners started differently as they showed many more P-forms than Group 2 learners at the beginning of the academic year. However, the learners in Group 1 equalled out by the end of the observation period. At the end of the academic year, the P forms ratio of the groups are similar.

We also counted types of pidginization features and found that the groups produced a rather similar percentage in each feature. L1 forms and constructions made up the majority of the features, followed by a lack of inflectional morphology and a lack of auxiliary. Instances of L1 mix were found mostly in the earlier sessions. Basic pidgin negation, lack of possessive inflection, and a lack of inversion in questions were also found, but not as frequently as the other features.

Together the findings suggest that learners make almost all progress in the first six months and then they stabilize in the forms and expressions that they use, which may be considered a fossilized system with typical pidginization features. Apparently, as the learners feel that they have a repertoire sufficient to communicate with each other, they do not make much progress anymore (cf. Schumann, 1978). During their interaction, the NTL output they produced was rarely corrected, probably because the learners had no clue that the forms were not target-like. Moreover, they were not asked to teach each other, but to use the L2 to communicate with. What we did see though is that Group 2 learners were more fluent as they took many more interactional turns than their Group 1 counterparts.

It was clear that the learners in the pesantren have only limited exposure to authentic or expert L2 input as the input they receive is mainly from their peers. Moreover, the type of instruction they receive from their teachers is mainly lexically based. The most common form of instruction is by means of a list of words to be used in their daily life. These factors may cause the learners progress to stagnate, as the developmental part of this study suggested. Finally, in terms of pidginization, the findings of Study 4 also suggest a role for the extensive peer interaction in promoting pidginization process. However, the suggested implication of extensive peer interaction on the L2 development still need to be interpreted with caution. Further studies in this topic in the context of pesantren still need to be done since there might be other factors contributing to the stagnation in the learners' English development.

Taking a DUB perspective, which holds that what is heard and used most frequently is what gets settled, the findings are not surprising as the learners only heard each other and were not really exposed to target forms. From a DUB perspective the frequency of use of L2 in social interaction, and the interaction of constructions in the network in the learner's mind (Verspoor & Behrens, 2011; Roehr-Brackin, 2015) is what drives the system. There are no innate systems, so the language learner can only discover and acquire the language through exposure and experience.

Together the findings of these studies suggest that when learners are asked to learn the L2 primarily through peer interaction with a list of given words, they may very well create their own language that suits them well, but it is not target like and may not be understood by speakers from other L2 English groups.

5.1. Limitations

There are numerous limitations in this study. First, this study was limited by the absence of the researcher in the writing sessions. During this process, the researcher was assisted by an English teacher in the pesantren who has been extremely helpful. However, this means the researcher cannot directly observe what happened on site. Moreover, since this process was done remotely in which the teachers needed to scan and upload the writings one by one, some learners' writings were missing leading to the omission of some learners which reduced the number of participants in this study. Furthermore, when administering the LHQ, learners had problems understanding technical terms in the questionnaire due to the learners' age although it had been translated to Indonesian. The learners needed to be guided through each question in the questionnaire, which took a lot

of effort. Therefore, it is suggested that a new language history questionnaire should be designed exclusively for young learners.

Another limitation of this study is the small sample size in Study 1 on peer interaction and Study 4 on pidginization. A larger sample size could have given a better idea of the interactional turns and pidginization features in the pesantren. However, this also would have required a greater amount of time as analyzing for specific features is very time consuming. Also, more actual classroom interaction between learner-learner and learner-teacher could have provided an interesting and deeper insight into the interactional practice in a pesantren. However, these limitations should be looked at as an opportunity for future research.

5.2. Implications

The findings of this study have a number of important implications for research and pedagogical practice. As far as interactionist research is concerned, this study shows how important it is to study interaction in real classroom settings and to study effects longitudinally. Almost all of the studies referred to in the background literature have studied interaction in very specific classroom contexts or laboratory settings in one session or perhaps a week and have considered uptake (e.g., Mackey and Goo, 2007; Mackey 2012; Loewen 2015) of a corrected form as "learning", but our study shows that in peer-to-peer interaction true communication seems to be the goal, and there is no corrective feedback so the learners copy each other's non target forms. Moreover, as far as we know, this has been the first longitudinal study of an interactionist approach to L2 development and it shows that the learners settle rather soon in a repertoire that is sufficient for them.

As far as pedagogical implications are concerned, it is remarkable what pesantrens have done in an effort to enable the learners to speak English, especially to conduct such practice in a country where English is a foreign language. They prove that learners do not have to live in a country where English is spoken as L1 to be able to communicate in English in daily life. It is amazing how the learners can memorize all the words given to them and use them effectively in communication. The theories that have inspired the instructional modes in the pesantren were set up many years ago when lexical theory and interaction theory were clearly on the rise, and comprehensible input had become less popular. Of course, at the time usage-based linguistic theories with their emphasis on exposure of whole utterances and chunks from which language patterns can be learned

inductively were not known yet in the field of SLA and might have informed instructional practice.

We feel that without changing the whole instructional approach too much, a few specific improvements could be made to help learners master the target forms better and to expand their repertoires. The main idea would be to put the words that the learners are to learn in short target like sentences, preferably in the form of target-like conversations. These conversations should especially contain target like chunks of language (such as the word *boarding school* instead of *boarding*). And as the learners early on do not have too much L2 yet, they should be asked to repeat, memorize and practice these short conversations, rather than making up their own creative sentences, so as to avoid the use of too many non-target-like forms during the first six weeks or so. Moreover, the finding of Study 3 particularly shows that the learners develop rapidly in the first six months. Therefore, in this critical period more attention should be given during the *mufradat/*vocabulary sessions and other English sessions to as much exposure to comprehensible authentic input as possible. Perhaps, in addition to the traditional lessons, students might see short videos (cf. Huang et al., 2021), or hear and read short stories (cf. Rousse-Malpat et al., 2019) where English is used in an actual authentic context so that they can expand their vocabulary and especially their use of chunks.

Also, the disadvantages of peer interaction can be minimized through some interventions. In a classroom-setting experiment by Sato and Lyster (2012) learners were trained on how to notice errors and to give feedback prior to interaction. This was done to minimize infrequent, inaccurate and unfocused feedback that is common in peer interaction. The results show that this intervention improved grammatical accuracy in learners' production. Sato and Lyster (2012) emphasized that learners need to realize that feedback is beneficial for both the provider and the receiver. A more recent study by Dao (2020) also suggested that interaction strategies need to be instructed to learners prior to interaction tasks. The results of Dao's study showed that the interaction strategy instruction generates more idea units, LREs, talk encouragement and reflection, as well as positive emotions among the participants. To sum up, more effort is needed to ensure that the learners L2 do not pidginize.

5.3. Future directions

This study has provided some insights on the effect of extensive learner-learner interaction on L2 development and pidginizaton particularly in the context of a pesantren.

This study also shows that a pesantren can provide a natural context for research in second language acquisition. However, several questions still remain to be answered. Since there are six different levels (Grade 7-12) in the pesantren, do all learners in the pesantren develop the same as the ones in this study? How do learners in the pesantren differ from learners in different education systems (e.g., public schools) in terms of their English development? Many other such questions require further investigation. More research using controll groups will also need to be done to further determine the effect of peer interaction in the context of the pesantren. Moreover, it would be interesting to document in detail what happens in their classrooms. Finally, it would be intriguing to see how the solutions suggested in this study (e.g., more authentic input) could benefit to the development of the learners. Finally, since the learners' development in the current study was operationalized using written texts, it would be interesting to see when learners' oral language is used instead.

REFERENCES

Adams, R. (2007). Do second language learners benefit from interacting with each other. *Conversational interaction in second language acquisition*, 29-51.

Adams, R., Nuevo, A. M., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner–learner interactions? *The Modern Language Journal*, *95*, 42-63.

Ahmad, D. M. (2012). The dynamics of the pondok pesantren: An Islamic educational institution in Indonesia. *Currents of Encounter*, 43, 63-74.

Al-Baekani, A. K. & Pahlevi, M. R. (2018). Investigation of English learning model at traditional Islamic boarding school (pondok pesantren salafiyah) Darul Ulum Al-Barokah Karawang. *Jurnal Pendidikan Unsika*, 6(2), 1-5.

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28(04). doi:10.1017/s0272263106060268

Andersen, R. (1981). Two perspectives on pidginization as second language acquisition. In R. W. Andersen (ed.), *New Dimensions in Second Language Acquisition Research.* Rowley, MA: Newbury House.

Anderson, E. A. (1997). The use of speech levels in Sundanese. In Clark, M. (ed.). *Papers in Southeast Asian Linguistics No. 16* (pp. 1–45)**.** Canberra: Pacific Linguistics. doi:10.15144/PL-A90.1

Aziez, F. (2016). An analysis of interlanguage performed by students of an Islamic boarding school in Tasikmalaya. *ELT Perspective, Jurnal Pendidikan Bahasa Inggris*, 4(2), 102-122.

Badan Pusat Statistik (BPS). (2010). *Kewarganegaraan, suku bangsa, agama, dan bahasa sehari-hari penduduk Indonesia* [Nationalities, ethnic groups, religions, and everyday languages of Indonesian population]. Badan Pusat Statistik Indonesia. Retrieved from: https://www.bps.go.id on 25 July 2020.

Baker, P., & Mühlhäusler, P. (1990). From business to pidgin. *Journal of Asian Pacific Communication*, 1(1), 87-115.

Bakker, P. (1994). Pidgins. In Arends, J., Muijsken, P., & Smith, N. *Pidgins and creoles: an introduction.* John Benjamins, 26–39.

Baralt, M. & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-meditation and face-to-face communication. *Language Teaching Research*, 15(2), 201–229.

Barlow, M., & Kemmer, S. (2000). Usage-Based Models. Sanford, CA: CSLI Publications.

Basterrechea, M. & P. Garcıa Mayo (2013). Language-related episodes during collaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough & A. Mackey (eds.). *Second language interaction in diverse educational contexts* (pp. 25–43). Philadelphia: John Benjamins.

Bickerton, D. (1977). Pidginization and creolization: language acquisition and language universals. In A. Valdman (ed.), *Pidgin and Creole Linguistics*. Bloomington: Indiana University Press.

Bigelow, M., Delmas, R., Hansen, K. & Tarone, E. (2006). Literacy and the processing of oral recasts in SLA. *TESOL Quarterly*, 40(4), 665–689.

Bin Tahir, S. Z. (2015). Multilingual behavior of Pesantren IMMIM students in Makassar. *The Asian EFL Journal Professional Teaching Articles.* August 2015, Issue 86. pp. 45-64.

Bin Tahir, S. Z. (2016). Multilingual teaching and learning at pesantren schools in Indonesia. *Asian EFL Journal*, 98, 74-94.

Bin Tahir, S. Z., Atmowardoyo, H., Dollah S., & Rinantanti, Y. (2017). Multilingual instructional model of pesantren schools in Indonesia. *Journal of Language Teaching and Research*, 8(6), 1210-1216.

Bowles, M., Toth, P. & Adams, R. (2014). A comparison of L2-L2 and L2-heritage learner interactions in Spanish language classrooms. *Modern Language Journal*, 98(2), 497–517.

Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research.* 20, 436–458.

van Bruinessen, M. (1994). Pesantren and kitab kuning: Continuity and Change in a Tradition of Religious Learning. In Texts from *the lslands: Oral and Wiritten Traditions of Indonesia and the Malay World (Ethnologica Bernica, 4)*,edited by Wolfgang Marschall, pp. 121-145, Berne: University of Berne.

van Bruinessen, M. (2006). Divergent Paths from Gontor: Muslim Educational Reform and the Travails of Pluralism in Indonesia. In F. L. Bakker and J. S. Aritonang (Eds.). *On the Edge of Many Worlds [Festschrift Karel A. Steenbrink]* (pp. 191–201). Zoetermeer: Meinema. available at: http://www.hum.uu.nl/medewerkers/m.vanbruinessen/publications/Bruinessen_Divergent_paths_from_Gontor.pdf

Bulté, B. (2013). *The development of complexity in Second Language Acquisition: A dynamic sys- tem approach.* (Unpublished doctoral dissertation). University of Brussels.

Cadierno, T., & Eskildsen, S. W. (Eds.). (2015). *Usage-based perspectives on second language learning* (Vol. 30). Berlin: Walter de Gruyter

Cancino, E., Rosansky, E. and Schumann, J. (1974). Testing hypotheses about second language acquisition: the copula and negative in three subjects. *Working Papers on Bilingualism*, 88-96.

Cao, Y. (2014). A sociocognitive perspective on second language willingness to communicate. *TESOL Quarterly,* 48(4), 789–814.

Cao, Y. & J. Philp (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System,* 34(4), 480–493.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (ed.). *Individual differences and universals in language learning aptitude* (pp. 83-118). Rowley, MA: Newbury House.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry & C. W. Stansfield (eds.). *Language aptitude reconsidered* (pp. 11–29). Washington, DC: Center for Applied Linguistics.

Carroll, J. B. & S. Sapon (2002). *Modern language aptitude test (MLAT). Manual 2002 edition.* N. Bethesda, MD: Second Language Testing, Inc.

Caspi, T. (2010). *A dynamic perspective on second language development.* PhD dissertation, University of Groningen.

Center for Applied Linguistics. (2011). *Directory of foreign language immersion programs in U.S. schools.* Retrieved April 1, 2017, from http://webapp.cal.org/Immersion/.

Crowhurst, M. & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English,* 13(2), 101-109.

Crystal, D. (2010). *The Cambridge encyclopedia of language*. Cambridge: Cambridge University Press.

Csizér, K. (2017). Motivation in the L2 classroom. In S. Loewen & M. Sato (eds.). The Routledge handbook of instructed second language acquisition (pp.418–432). New York: Taylor & Francis

Cummins, J. (2009). *Bilingual and Immersion Programs.* The Handbook of Language Teaching. Malden, MA: Wiley-Blackwell.

Dao, P. (2020). Effect of interaction strategy instruction on learner engagement in peer interaction. *System*, 91, 102244.

Daulay, H. P. (2009). *Dinamika pendidikan Islam di asia tenggara [The dynamics of Islamic education in southeast Asia]*. Jakarta, Indonesia: Rineka Cipta.

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning,* 46(3), 529–555.

de Bot, K. (2008). Introduction: Second language development as a dynamic process. *The Modern Language Journal,* 92(2), 166–178.

de Bot, K. & Larsen-Freeman, D. (2011) Researching second language development from a dynamic systems theory perspective. In M. H. Verspoor, K. de Bot & W. Lowie (Eds.) *A Dynamic Approach to Second Language Development: Methods and Techniques* (pp. 5–23). Amsterdam: John Benjamins.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism*, *10*(1), 7.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, *11*(4), 227-268.

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian psychology/Psychologie canadienne*, *49*(1), 14.

Deci, E. L., & Ryan, R. M. (2014). Autonomy and need satisfaction in close relationships: Relationships motivation theory. In *Human motivation and interpersonal relationships* (pp. 53-73). Dordrecht: Springer

DeKeyser, R. M. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning,* 62(2), 189–200.

DeKeyser, R. M. (2017a). Knowledge and skill in ISLA. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.15–32). New York: Taylor & Francis

DeKeyser, R. M. (2017b). Age in learning and teaching grammar. In J. Liontas (ed.), *The TESOL encyclopedia of English language teaching.* New York: Wiley.

Dewaele, J. M. (2017). Personality and anxiety. In S. Loewen & M. Sato (eds.) *The Routledge handbook of instructed second language acquisition* (pp. 433–450). New York: Taylor & Francis

Dhofier, Z. (1985). *Tradisi Pesantren: studi tentang pandangan hidup kyai.* Jakarta: LP3ES.

Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (ed.). *Individual differences and instructed language learning* (pp. 137-158). Amsterdam: John Benjamins.

Dörnyei, Z. & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research,* 4(3), 275–300.

Doughty, C. & Pica, T. (1986). 'Information gap' tasks: An aid to second language acquisition? *TESOL Quarterly*, 20(2), 305–325.

Eberhard, D. M., Gary F. S., and Charles D. F. (eds.). (2020). *Ethnologue: Languages of the World. Twenty-third edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com. Retrieved 25 July 2020.

Eckerth, J. (2008). Investigating consciousness-raising tasks: pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics,* 18(2), 119–145.

Eckerth, J. (2009). Negotiated interaction in the L2 classroom. *Language Teaching,* 42(1), 109–130.

Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics.* Oxford: Blackwell, 525–551.

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications.* New York: Routledge, 3–18.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition,* 21(2), 285–301.

Ellis, R., H. Basturkmen & Loewen, S. (2001a). Learner uptake in communicative ESL lessons. *Language Learning,* 51(2), 281–318.

Ellis, R., H. Basturkmen & Loewen, S. (2001b). Preemptive focus on form in the ESL classroom. *TESOL Quarterly,* 35(3), 407–432.

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System,* 30(4), 419–432.

Ellis, R., Loewen, S., & Erlam, R. (2009). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Implicit and Explicit Knowledge in Second*

Language Learning, Testing and Teaching, 303–332. doi:10.21832/9781847691767-015

Farid, A., & Lamb, M. (2020). English for Da'wah? L2 motivation in Indonesian pesantren schools. *System,* 102310. doi: https://doi.org/10.1016/j.system.2020.102310

Ferguson, C. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, 17(1), 1-14.

Ferguson, C. (1977). Baby talk as simplified register. In C. E. Snow and C. Ferguson (eds), *Talking to Children: Language input and acquisition.* Cambridge: Cambridge University Press, 209-235.

Fernández Dobao, A. (2016). Peer interaction and learning: A focus on the silent learner. In Sato, M. & Ballinger, S. (eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda.* Amsterdam: John Benjamins, 33–61.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics,* 19(1), 1–23.

Freeman, Y. (2005). *Dual language essentials for teachers and administrators*. Heinemann: Portsmouth.

Garcıa Mayo, M. P. & Pica, T. (2000). L2 learner interaction in a foreign language setting: Are learning needs addressed? *International Review of Applied Linguistics,* 38(1), 35–58.

Garcıa Mayo, M. P. & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact on language-related episodes? In M. Sato & S. Ballinger (eds.). *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 241-266). Amsterdam: John Benjamins.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Edward Arnold.

Gass, S. M. (1997). *Input, interaction and the second language learner.* Mahwah, NJ: Lawrence Erlbaum.

Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. H. Long (eds.), *The handbook of second language acquisition.* Malden, MA: Blackwell Publishing, 224–255.

Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-Based Interactions in Classroom and Laboratory Settings. *Language Learning,* 55(4), 575–611.

Gass, S. M. & Mackey, A. (2000). *Stimulated recall methodology in second language research.* Mahwah, NJ: Lawrence Erlbaum.

Gass, S. M. & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (eds.), *Theories in second language acquisition* (2nd ed.). New York: Routledge, 180–206.

Gass, S. M., & Mackey, A. (2020). Input, Interaction, and Output in L2 Acquisition. *Theories in Second Language Acquisition*, 192–222. doi:10.4324/9780429503986-9

Gass, S. M. & Varonis, E. (1985). Variation in native speaker speech modification to non-native speakers. *Studies in Second Language Acquisition,* 7(1), 37–57.

Gass, S. M. & E. Varonis (1986). Sex differences in NNS/NNS interactions. In R. Day (ed.). *Talking to learn: Conversation in second language interaction* (pp. 327-351). Rowley, MA: Newbury House.

Gass, S. M. & Varonis, E. (1990). Miscommunication in nonnative speaker discourse. In Coupland N., Giles H. & Wiemann, J. (eds.), *'Miscommunication' and problematic talk*. London, New Delhi: Sage Publications, 121–145.

Gass, S., & Varonis, E. (1994). Input, Interaction, and Second Language Production. *Studies in Second Language Acquisition,* 16(3), 283-302.

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34(3), 445–474.

Goo, J. & Mackey, M.. (2013). The case against the case against recasts. *Studies in Second Language Acquisition,* 35(1), 1–39.

Hammerly, H. (1991). *Fluency and Accuracy: Toward balance in language teaching and learning.* Clevedon: Multilingual Matters.

Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (ed.), *Second language acquisition: A book of readings*. Rowley, MA: Newbury House, 401–435.

Hawkins, B. (1985). Is an "appropriate response" always so appropriate? In S. Gass &C. Madden (Eds.), *Input in second language acquisition* (pp. 162–178). Rowley, MA: Newbury House.

Hidayat, M. (2007). *The Teaching of English at Pondok Pesantren At Tarbiyyah Al Islamiyyah in Paiton.* Probolinggo: The State Islamic University of Malang.

Hong, N. (2013). *Dynamic usage-based approach to second language teaching.* (Unpublished PhD dissertation). Rijksuniversiteit Groningen, Groningen.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70, 125-132.

Hou, J., Verspoor, M. H., & Loerts, H. (2016). An exploratory study into the dynamics of Chinese L2 writing development. *Dutch Journal of Applied Linguistics,* 5(1), 65–96.

Huang, T., Steinkrauss, R., & Verspoor, M. (2021). Variability as predictor in L2 writing proficiency. *Journal of Second Language Writing,* doi: https://doi.org/10.1016/j.jslw.2020.100787

Indonesian Ministry of Religious Affairs. (2020). *Statistik pesantren.* Retrieved from https://ditpdpontren.kemenag.go.id/pdpp

Irshad, M. (2015). *A dynamic approach to autonomous second language development*. (Unpublished PhD dissertation). Rijksuniversiteit Groningen, The Netherlands.

Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System*, 29(2), 267–287.

Jihad, A. (2011). *Efektifitas hukuman terhadap kedisiplinan santri di Pondok Pesantren Daar el-Qolam.* (Bachelor thesis, UIN Syarif Hidayatullah, Indonesia). Retrieved from http://repository.uinjkt.ac.id/dspace/bitstream/123456789/1439 /1/100887-AKHMAD JIHAD-FITK.pdf

Johnson K. & Johnson, H. (1999). *Encyclopedic dictionary of applied linguistics: a handbook for language teaching.* Oxford: Blackwell Publishers

Jubaidah, S. (2015). Pembelajaran dan pemerolehan bahasa Arab di pondok modern Gontor di Darul Ma'rifat Gurah Kediri Jatim. *Jurnal Parameter*, 27(2). doi: doi.org/10.21009/parameter.272.09

Kang, K. I. (2015). Peer interaction: a compromise or a necessity? *APPLE award winning papers in TESOL & AL*, 15(2).

Kim, Y. (2017). Cognitive-interactionist approaches to L2 instruction. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.126–135). New York: Taylor & Francis

Kim, Y. & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12(2), 211–234.

Kim, Y., Payant, C. & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory. *Studies in Second Language Acquisition.* 37(3), 549–581.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kenner (Eds.), *Usage- based models of language* (pp. 1–63). Stanford: CSLI Publications.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Langacker, R. W. (2009). A dynamic view of usage and language acquisition. *Cognitive Linguistics, 20*(3), 627-640.

Larsen-Freeman, D. E. (1976). An explanation for the morpheme acquisition order of second language learners. *Language learning*, *26*(1), 125-134.

Larsen-Freeman, D. (1997). Chaos/complexity science and Second Language Acquisition. *Applied Linguistics*, 18(2), 140–165.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics*, *27*(4), 590-619.

Larsen-Freeman, D. (2006). Second language acquisition and the issue of fossilization: There is no end, and there is no state. In Z. Han & T. Odlin (Eds.), *Studies of fossilization in second language acquisition* (pp. 189-200). Clevedon, England: Multilingual Matters.

Larsen-Freeman, D. & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. *The modern language journal*, *92*(2), 200-213.

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25(1), 37–63.

Lefebvre, C. (2004). *Issues in the study of pidgins and creole languages.* Amsterdam: John Benjamins B.V.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365.

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal,* 97(3), 634–654.

Li, S. (2017). Cognitive differences and ISLA. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.396–417). New York: Taylor & Francis

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism, 17*(3), 673.

Li, S., Zhu, Y. & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structures. *Modern Language Journal*. 100(1), 276–295.

Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (eds.). *Focus on form in classroom second language acquisition* (pp. 177–196). Cambridge: Cambridge University Press.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning,* 54(1), 153–187.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition,* 27(3), 361–386.

Loewen, S. (2015). *Introduction to instructed second language acquisition.* New York: Routledge.

Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukasa, K., Ahn, S. & Chen, X. (2009). Second language learners' beliefs about grammar instruction and error correction. *Modern Language Journal*, 93(1), 91–104.

Loewen, S. & Nabei, T. (2007). *Measuring the effects of oral corrective feedback on L2 knowledge.* In A. Mackey (ed.), 361–377.

Loewen, S. & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal.* 90(4), 536–556.

Loewen, S. & Sato, M. (2017). Instructed second language acquisition (ISLA): An overview. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.1–12). New York: Taylor & Francis

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329. doi:10.1017/s0261444818000125

Loewen, S. & Wolff, D. (2016). Peer interaction in F2F and CMC contexts. In M. Sato & S. Ballinger (eds.). *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp 163–184). Amsterdam: John Benjamins.

Long, M. H. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences* 379, 259–278.

Long, M. H. (1983a). Linguistic and conversational adjustments to nonnative speakers. *Studies in Second Language Acquisition*, 5, 177–194.

Long, M. H. (1983b). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126–141.

Long, M. H. (1985). Input and second language acquisition theory. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 377–393). Rowley, MA: Newbury House.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), Handbook of second language acquisition (pp. 413–468). New York: Academic Press.

Long, M. H. (2007). *Problems in SLA.* New York: Routledge.

Long, M. H. (2015). *Second language acquisition and task-based language teaching.* New York: Wiley.

Long, M. H. (2017). Instructed second language acquisition (ISLA): Geopolitics, methodological issues, and some major research questions. *Instructed Second Language Acquisition* 1, 7–44.

Long, M. H. & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly* 19.2, 207–228.

Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82(3), 357–371.

Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition*, 16(3), 303–323.

Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, 69, 184-206.

Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48(2), 183–218.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition,* 26(3), 399–432.

Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66.

Lyster, R. & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, 35(1), 1–18.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom: A meta-analysis. *Studies in Second Language Acquisition,* 32(2), 265–302.

Lyster, R. & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. Garcıa Mayo, J. Gutierrez-Mangado & M. Mart´ınez Adrian (eds.), *Contemporary approaches to second language acquisition.* Amsterdam: John Benjamins, 71–92.

Lyster, R., Saito, K. & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40.

MacIntyre, P. (1995). How does anxiety affect foreign language learning? A reply to Sparks and Ganschow. *The Modern Language Journal*, 79, 90-99.

MacIntyre, P. D., Burns, C. & Jessome, A. (2011). Ambivalence about communicating in a second language: A qualitative study of French immersion students' willingness to communicate. *Modern Language Journal*, 95(1), 81–96.

MacIntyre, P. D., Z. Dörnyei, R. Clément & K. A. Noels (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, 82(4), 545–562.

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21(4), 557–587.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. Applied Linguistics 27(3), 405–430.

Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning.* Oxford: Oxford University Press.

Mackey, A., Abbuhl, R. & Gass, S. M. (2012). Interactionist approach. In Gass, S. M. & Mackey, A. (eds.) *The Routledge handbook of second language acquisition.* New York: Routledge, 7–24.

Mackey, A. & Gass, S. M. (2015). *Second language research: Methodology and design.* New York: Routledge.

Mackey, A., Gass, S. M. & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471–497.

Mackey, A. & Goo, J. (2007). *Interaction research in SLA: A meta-analysis and research synthesis.* In A. Mackey (ed.), 407–449.

Mackey, A., Kanganas, A. P. & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41(2), 285–312.

Mackey, A., Oliver, R. & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning*, 53(1), 35–66.

Mackey, A., Fujii, A., Biesenbach-Lucas, S., Weger, H., Jacobsen, N., Fogle, L., Lake, J., Sondermann, K., Tagarelli, K., Tsujita, M., Watanabe, A., Abbuhl, R. & Kim, K. (2013). Tasks and traditional practice activities in a foreign language context. In K. McDonough & A. Mackey (eds.), 71–87.

Master, P., Schumann, J., and Sokolik, M. E. (1989). The experimental creation of a pidgin language. *Journal of Pidgin and Creole Languages*, 4(1), 37–63.

McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32(2), 207–224.

McDonough, K. (2017). Experimental research methods. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.562–576). New York: Taylor & Francis

McDonough, K., W. J. Crawford & A. Mackey (2015). Creativity and EFL students' language use during a group problem-solving task. *TESOL Quarterly*, 49(1), 188–199.

McLaughlin, B. (1987). *Theories of Second-Language Learning.* London: Edward Arnold.

Mühlhäusler, P. (1986). *Pidgin and Creole Linguistics.* Oxford: Blackwell.

Nakatsukasa, K. (2016). Efficacy of recasts and gestures on the acquisition of locative prepositions. *Studies in Second Language Acquisition*, 38(4), 771–799.

Nassaji, H. (2013). Participation structure and incidental focus on form in adult ESL classrooms. *Language Learning*, 63(4), 835–869.

Nemser, W. (1974). Approximative systems of foreign language learners. In J. C. Richards (ed.), *Error Analysis: Perspectives on second language acquisition.* London: Longman, 55-63.

Oliver, R. (1998). Negotiation of meaning in child interactions. The relationship between conversational interaction and second language acquisition. *Modern Language Journal*, 82(3), 372–386.

Oliver, R. (2000). Age difference in negotiation and feedback in classroom and pair work. *Language Learning,* 50(1), 119–151.

Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *Modern Language Journal,* 86(1), 97–111.

Oliver, R. & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *Modern Language Journal,* 87(4), 519–533.

Oliver, R., Nguyen, B. & Sato, M. (2017). Child instructed SLA In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp. 468– 487). New York: Taylor & Francis.

Oliver, R., Philp, J. & Duchesne, S. (2017). Children working it out together: A comparison of younger and older learners collaborating in task based interaction. *System,* 69, 1–14.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics*, *24*(4), 492-518.

Oxford, R.L. (1993). Instructional implications of gender differences in language learning styles and strategies. *Applied Language Learning* 4, 65-94.

Oxford, R. L., M. Nyikos & M. E. Ehrman (1988). Vive la différence? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals* 21, 321- 329.

Panggabean, H. (2015). Problematic Approach to English Learning and Teaching: A Case in Indonesia. English Language Teaching, 8(3), 35-45

Pérez-Vidal, C. (2017). Study abroad and ISLA. In S. Loewen & M. Sato (eds.). *The Routledge handbook of instructed second language acquisition* (pp.339–360). New York: Taylor & Francis

Philp, J. (2003). Constraints on 'noticing the gap': Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition,* 25(1), 99–126.

Philp, J., Adams, R. & Iwashita, N. (2014). *Peer interaction and second language learning.* New York: Routledge.

Pica, T. (1988). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning,* 38(1), 45–73.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning,* 44(3), 493–527.

Pica, T. (2013). From input, output and comprehension to negotiation, evidence, and attention: An overview of theory and research on learner interaction and SLA. In M. Garcıa Mayo, J. Gutierrez-Mangado & M. M. Adrian (eds.), *Contemporary approaches to second language acquisition*. Amsterdam: John Benjamins, 49–69.

Pica, T., Young, R., & Doughty, C. (1987). The Impact of Interaction on Comprehension. *TESOL Quarterly*, 21(4), 737.

Pica, T., L. Holliday, N. Lewis, D. Berducci & J. Newman. (1991). Language learning through interaction. *Studies in Second Language Acquisition,* 13(2), 343–376.

Pica, T., F. Lincoln-Porter, D. Paninos & J. Linnell (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly,* 30(1), 59–84.

Plonsky, L. & S. M. Gass (2011). Quantitative research methods, study quality and outcomes: The case of interaction research. *Language Learning,* 61(2), 325–366.

Plonsky, L. & D. Brown. (2015). Domain definition and search techniques in meta-analyses of L2 research (or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31(2), 267–278.

Poerbakawaba, S. (1976). *Ensiklopedia pendidikan*. Jakarta: Gunung Agung

Porter, P. (1986). How learners talk to each other: Input and interaction in task-centered discussions. In R. Day (ed.), *Talking to learn: Conversation in second language acquisition.* Rowley, MA: Newbury House, 200–222.

Qomar, M. (2005). *Pesantren: dari transformasi metodologi menuju demokrasi institusi.* Jakarta: Erlangga

Raswan. (2017). Tamyiz; model alternatif pembelajaran bahasa Arab sebagai bahasa Al-Qur'an. *Lisanul Arab, Journal of Arabic Learning and Teaching*, 6(1), 18-28.

Rassaei, E. (2015). The effects of foreign language anxiety on EFL learners' perceptions of oral corrective feedback. *Innovation in Language Learning and Teaching*, 9(2), 87–101.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *Modern Language Journal* 95(2), 162–181.

Révész, A. (2012). Working memory and the observed effectiveness of recasts on different outcome measures. *Language Learning*, 62(1), 93–132.

Richards, J. C. (1974). Social factors, interlanguage, and language learning. In J. C. Richards (ed.), *Error Analysis: Perspectives on second language acquisition*. London: Longman, 64-91. [Originally published in 1972 in Language Learning, 22(2), 159-188.]

Richards, J. C & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics.* Harlow: Pearson Education.

Risdianto, F. (2016). Model pembelajaran bahasa Inggris di ponpes 'Ta'mirul Islam' Surakarta. *Leksema, Jurnal Bahasa dan Sastra,* 1(1), 47-58.

Ritonga, T., Ananda, A., Lanin, D.,& Hasan, H. (2019). Practice to theory of learning: a lesson learned from Islamic boarding school in South Tapanuli of Indonesia. *Humanities & Social Sciences Reviews,* 5(2019), 1304-1310.

Robinson, P. (1995). Attention, memory and the 'noticing' hypothesis. *Language Learning*, 45(2), 99–140.

Robinson, P. (1996). *Consciousness, rules and instructed second language acquisition.* New York: Lang.

Robinson, P. (2003). *Attention and memory during SLA.* In Doughty, C. J. & Long, M. H. (eds.), 631–678.

Roehr-Brackin, K. (2015). Long-term development in an instructed adult L2 learner: Usage-based and complexity theory applied. *Usage-based perspectives on second language learning*, 181-206.

Romaine, S. (1988). *Pidgin and Creole Languages.* London: Longman.

Rousse-Malpat, A., Steinkrauss, R., & Verspoor, M. (2019). Structure-based or dynamic usage?based instruction: long-term effects on (morpho)syntactic and lexical complexity in writing samples. *Instructed Second Language Acquisition*, 3(2), 181–205. https://doi.org/10.1558/isla.38054

Russell, J. & N. Spada. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (eds.), *Synthesizing research on language learning and teaching.* Philadelphia: John Benjamins, 133–164.

Ryan, R. M., & Deci, E. L. (2008). A self-determination theory approach to psychotherapy: The motivational basis for effective change. *Canadian Psychology/Psychologie canadienne*, *49*(3), 186.

Saito, K., Suzuki, S., Oyama, T., & Akiyama, Y. (2019). How does longitudinal interaction promote second language speech learning? Roles of learner experience and proficiency levels. *Second Language Research*.

Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal,* 97(3), 611–633.

Sato, M. (2015). Density and complexity of oral production in interaction: The interactionist approach and an alternative. *International Review of Applied Linguistics in Language Teaching*, 53(3), 307–329.

Sato, M. (2016). Interaction mindsets, interactional behaviors, and L2 development: An affective- social-cognitive model. *Language Learning*, 67(2).

Sato, M. & R. Lyster (2007). *Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types.* In A. Mackey (ed.), 123–142.

Sato, M. & R. Lyster (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition,* 34(4), 591–262.

Sato, M. & Viveros, P. (2016). Interaction or collaboration? Group dynamics in the foreign language classroom. In M. Sato & S. Ballinger (eds.). *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp 91-112). Amsterdam: John Benjamins.

Sauro, S. & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics*, 31(4), 554–577.

Saville-Troike, M. & Barto, K. (2016). Introducing second language acquisition (3rd Edition). Cambridge: Cambridge University Press.

Schachter, J. (1991). Corrective feedback in historical perspective. *Second Language Research*, 7(2), 89–102.

Schegloff, E. (2000). When 'others' initiate repair. *Applied Linguistic*s, 21(2), 205–243.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11, 129–158.

Schmidt, R. (1995). *Attention and awareness in foreign language learning.* Honolulu: University of Hawaii Press.

Schmidt, R. (2001). Attention. In P. Robinson (ed.), *Cognition and second language instruction.* Cambridge: Cambridge University Press, 3–32.

Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing,* 22,1–30.

Schoonen, R. (2012). The validity and generalizability of writing scores: The effect of rater, task and language. In E. Van Steendam (Ed.), *Measuring writing: Recent insights into theory, methodology and practices* (pp. 1–22). Leiden: Koninklije Brill.

Schulz, R. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback. *Modern Language Journal,* 85, 244–258.

Schumann, J. H. (1974). The implication of interlanguage, pidginization and creolization for the study of adult second language acquisition. *TESOL Quarterly,* 8(2), 145-152.

Schumann, J. H. (1978). *The pidginization process: a model for second language acquisition.* Rowley, MA: Newbury House.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics,* 10, 209-231.

Selinker, L. (1992). *Rediscovering interlanguage.* London: Longman.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin Company.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263–300.

Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning,* 58(4), 835–874.

Sheen, Y. & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning, vol. 2.* New York: Routledge, 593–610.

Shehadeh, A. (1999). Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning,* 49(4), 627–675.

Shehadeh, A. (2001). Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly,* 35(3), 433–457.

Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

Sneddon, J. (2003). *The Indonesian Language: Its history and role in modern society.* Sydney: University of South Wales Press.

Sofwan, M., & Habibi, A. (2016). Problematika dunia pendidikan Islam abad 21 dan tantangan pondok Pesantren di Jambi [Problems in the world of 21st century Islamic education and the challenges of Islamic boarding schools in Jambi]. *Jurnal Kependidikan [Education Journal]*, 46(2), 271-280.

Spada, N. & P. Lightbown. (2009). Interaction research in second/foreign language classrooms. In A. Mackey & C. Polio (eds.), *Multiple perspectives on interaction.* New York: Routledge, 157–175.

Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, *31*(4), 532-553.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158.

Storch, N. & A. Aldosari (2013). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31–48.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (eds.), *Input in second language acquisition.* Rowley, MA: Newbury, 235–253.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (eds.), *Principle and practice in applied linguistics*. Oxford: Oxford University Press, 125–144.

Swain, M. (1997). French immersion programs in Canada. In: Cummins J., Corson D. (eds) *Bilingual education. Encyclopedia of language and education, vol 5*. Dordrecht: Springer.

Swain, M. (2005). *The output hypothesis: Theory and research.* In E. Hinkel (ed.), 471–483.

Swain, M. & S. Lapkin (1998). Interaction and second language learning: two adolescent French immersion students working together. *Modern Language Journal* 82(3), 320–337.

Todd, L. (1974). *Pidgins and Creoles.* London: Routledge and Kegan Paul.

Tomasello, M. (2003). *Constructing language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tomita, Y. & Spada, N. (2013). Form-focused instruction and learner investment in L2 communication. *Modern Language Journal*, 97(3), 591–610.

Trofimovich, P., Ammar, A. & Gatbonton, E. (2007). How effective are recasts: The role of attention, memory, and analytical ability. In A. Mackey (ed.), 171–195.

Tyler, A. E., Ortega, L., Uno, M., & Park, H. I. (Eds.). (2018). *Usage-inspired L2 instruction: Researched pedagogy (Vol. 49).* Amsterdam: John Benjamins Publishing Company.

Umar, N. (2014). *Rethinking pesantren*. Jakarta: PT Elex Media Komputindo.

Utvær, B. K. S., & Haugan, G. (2016). The academic motivation scale: dimensionality, reliability, and construct validity among vocational students. *Nordic Journal of Vocational Education and Training*, *6*(2), 17-45.

Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological review*, *98*(1), 3.

VanPatten, B. (2017). Situation instructed language acquisition: Facts about second language acquisition. *Instructed Second Language Acquisition,* 1(1), 45–60.

Varonis, E. & S. M. Gass (1985). Miscommunication in native/non-native conversation. *Language in society*, 14(3), 327–343.

Verspoor, M. (2017). Complex Dynamic Systems Theory and L2 pedagogy. Lessons to be learned. In L. Ortega and Z. Han (Eds.), *Complexity Theory and Language Development* (pp. 143-162). Amsterdam: John Benjamins.

Verspoor, M. H., & Behrens, H. (2011). Dynamic systems theory and a usage based approach to second language development. In M. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: methods and techniques* (pp. 25 - 38). Amsterdam: John Benjamins.

Verspoor, M. & de Bot, K. (2021). Measures of variability in transitional phases in second language development. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2021-0026

Verspoor, M.H., de Bot, K., & Lowie, W. (Eds.) (2011). *A dynamic approach to second language development: Methods and techniques*. Amsterdam: John Benjamins.

Verspoor, M.H., de Bot, K., & Xu, W. (2011). The role of input and scholastic aptitude in second language development. *Toegepaste Taalwetenschap in Artikelen*, 86(1), 47 – 60.

Verspoor, M., de Bot, K., & Xu, X. (2015). The effects of English bilingual education in the Netherlands. *Journal of Immersion and Content-Based Language Education*, 3(1), 4–27. doi:10.1075/jicb.3.1.01ver

Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal, 92*(2), 214-231.

Verspoor, M.H., Schmid, M.S., & Xu, X. (2012) A dynamic usage-based perspective on L2 writing. *Journal of Second Language Writing*, 21, 239–263.

Verspoor, M., & Smiskova, H. (2012). Foreign language writing development from a dynamic usage based perspective. *L2 writing development: Multiple perspectives*, 17-46.

Vries, F. D., & Verspoor, M. (2010). Fouten maken mag in tweede taalontwikkeling!. *Levende Talen Tijdschrift*, 11(3), 18-28.

Wagner-Gough, J. & Hatch, E. (1975). The importance of input data in second language acquisition studies. *Language Learning* 25(2), 297–307.

Watanabe, Y. & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121–142.

Wekke, I. S., & Hamid, S. (2013). Technology on Language Teaching and Learning: A Research on Indonesian Pesantren. *Procedia - Social and Behavioral Sciences, 83, 585–589.* doi: 10.1016/j.sbspro.2013.06.111

Wekke, I. S. (2015).Arabic teaching and learning: a model from Indonesian Muslim minority. *Procedia - Social and Behavioral Sciences*, 191(2015), 286 – 290.

Williams, J. (2001a). Learner-generated attention to form. *Language Learning* 51(1), 303–346.

Williams, J. (2001b). The effectiveness of spontaneous attention to form. *System* 29(3), 325–340.

Xu, J., Fan, Y., & Xu, Q. (2019). EFL learners' corrective feedback decision-making in task-based peer interaction. *Language Awareness*, 28(4), 329–347.

Yang, Y. & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235–263.

Young, D. J. & R. L. Oxford (1997). A gender-related analysis of strategies used to process written input in the native language and a foreign language. *Applied Language Learning* 8, 1-20.

Yoshida, R. (2013). Conflict between learners' beliefs and actions: speaking in the classroom. *Language Awareness,* 22(4), 371–388.

Yule, G. & Macdonald, D. (1990). Resolving referential conflicts in L2 interaction: The effect of proficiency and interactive role. *Language Learning,* 40(4), 539–556.

Zakaria, G. A. N. (2010). Pondok pesantren: changes and its future. *Journal of Islamic and Arabic Education* 2(2): 45-52.

Zein, S. (2020). Language policy in superdiverse Indonesia. New York and London: Routledge.

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition.* 38(3), 553–586.

Ziemek, M. (1986). Pesantren dalam perubahan sosial. Translated by Butje B. Soendjojo from *Pesantren Islamisce bildung in Sozialen Wandel*. Jakarta: P3M

APPENDICES

Appendix A. Language History Questionnaire

L2 Language History Questionnaire (Version 3.0, 2015)

See http://blclab.org/ for online use and credit

Participant ID:

1. Usia (dalam tahun): _____

2. Jenis kelamin (Lingkari salah satu): Laki-laki/Perempuan

3. Pendidikan (saat ini atau pendidikan terakhir) (Lingkari salah satu):

- Pascasarjana (S3)
- Pascasarjana (S2)
- Sarjana (S1/D3/D1)
- SMA/MA kelas _____
- SMP/MTs kelas _____
- Lainnya (Sebutkan): _____

4. Pernahkan Anda mempelajari bahasa lain dalam hal menyimak, berbicara, membaca, atau menulis? (Lingkari salah satu)

Ya/Tidak

5. Sebutkan bahasa Ibu Anda dan bahasa lain yang pernah Anda pelajari, usia saat pertama Anda mempelajari bahasa tersebut dalam hal menyimak, berbicara, membaca, atau menulis, dan total jumlah tahun Anda mempelajari masing-masing bahasa tersebut.

Bahasa	Menyimak	Berbicara	Membaca	Menulis	Jumlah tahun penggunaan[a]

a. Anda mungkin pernah mempelajari satu bahasa, kemudian berhenti, dan kemudian mempelajarinya lagi. SIlahkan berikan total jumlah tahun Anda mempelajari bahasa tersebut.

6a. Negara/daerah tempat tinggal: _____

6b. Negara/daerah asal: _____

6c. Jika 6a dan 6b berbeda, kapan Anda pertama pindah ke negara/daerah Anda tinggal saat ini? _____

7. Jika Anda pernah tinggal atau mengunjungi negara/daerah selain negara/daerah tempat tinggal saat ini atau negara/daerah asal selama tiga bulan atau lebih, maka sebutkan nama negara/daerah tersebut, bahasa yang anda gunakan di sana, dan frekuensi penggunaan bahasa tersebut untuk masing-masing negara/daerah.

Negara/Daerah	Lama tinggal[a] (dalam bulan)	Bahasa	Frekuensi penggunaan[b]
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7

a. Anda mungkin pernah ke suatu negara/daerah tertentu beberapa kali untuk masa yang berbeda-beda. Jumlahkan total lama kunjungan Anda.

b. Silahkan beri nilai sesuai dengan skala berikut (lingkari angka dalam tabel di atas)

Tidak pernah	Jarang	Kadang-kadang	Rutin	Sering	Terbiasa	Selalu
1	2	3	4	5	6	7

8. Sebutkan usia Anda saat pertama kali Anda mengunakan bahasa-bahasa yang pernah Anda pelajari di lingkungan berikut.

Bahasa	Di rumah	Dengan teman	Di sekolah	Di tempat kerja	Software bahasa	Online games

9. Sebutkan bahasa yang digunakan oleh guru Anda di kelas di tiap tingkatan pendidikan. Jika bahasa di kelas pernah beralih dalam tingkat pendidikan tertentu, maka sebutkan juga “beralih ke” bahasa apa.

	Bahasa	(Beralih ke)
Sekolah dasar		
SMP/MTs		
SMA/MA		
Perguruan tinggi/universitas		

10. Beri nilai terhadap kemampuan belajar bahasa Anda. Dengan kata lain, seberapa baik diri Anda, menurut pribadi, dalam mempelajari bahasa baru, dibanding dengan teman-teman atau orang lain yang Anda kenal? (Lingkari salah satu)

Sangat lemah	Lemah	Terbatas	Rata-rata	Baik	Sangat Baik	Unggul
1	2	3	4	5	6	7

11. Beri nilai kemampuan bahasa Anda saat ini dalam kemampuan menyimak, berbicara, membaca, dan menulis dalam tiap-tiap bahasa yang pernah Anda pelajari. Silahkan nilai berdasarkan skala berikut (lingkari angka dalam table):

Sangat lemah	Lemah	Terbatas	Fungsional	Baik	Sangat baik	Mirip penutur asli
1	2	3	4	5	6	7

Bahasa	Menyimak	Berbicara	Membaca	Menulis
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

12. Jika Anda pernah mengikuti tes kemampuan berbahasa yang standar (missal TOEFL), maka sebutkan nama tesnya, bahasa yang dinilai, dan skor yang Anda terima untuk masing-masing tes. Jika Anda tidak ingat skor pastinya, maka sebutkan "perkiraan skor" tes tersebut.

Tes	Bahasa	Skor	(Perkiraan skor)

13. Beri nilai tingkat aksen asing Anda dalam tiap-tiap bahasa yang pernah Anda pelajari. Beri nilai sesuai dengan skala berikut (lingkari angka dalam tabel):

Tidak ada	Sangat lemah	Lemah	Sedang	Kuat	Sangat kuat	Ekstrim
1	2	3	4	5	6	7

Bahasa	Kekentalan aksen
	1 2 3 4 5 6 7
	1 2 3 4 5 6 7
	1 2 3 4 5 6 7
	1 2 3 4 5 6 7

14. Perkirakan berapa jam per hari yang Anda habiskan untuk aktivitas berikut dalam tiap-tiap bahasa yang Anda pernah pelajari.

	Bahasa: _____	Bahasa: _____	Bahasa: _____
Menonton televisi:	_____(jam)	_____(jam)	_____(jam)
Mendengarkan radio/musik:	_____(jam)	_____(jam)	_____(jam)
Membaca santai:	_____(jam)	_____(jam)	_____(jam)
Membaca untuk pekerjaan/sekolah:	_____(jam)	_____(jam)	_____(jam)
Menulis di sosial media/email/surat kepada teman:	_____(jam)	_____(jam)	_____(jam)
Menulis untuk pekerjaan/sekolah:	_____(jam)	_____(jam)	_____(jam)

15. Perkirakan berapa jam per hari yang Anda habiskan untuk berbicara dalam kelompok orang berikut dalam bahasa yang pernah Anda pelajari.

	Bahasa: _____	Bahasa: _____	Bahasa: _____
Anggota keluarga:	_____(jam)	_____(jam)	_____(jam)
Teman[a]:	_____(jam)	_____(jam)	_____(jam)
Teman sekelas:	_____(jam)	_____(jam)	_____(jam)
Teman kerja[b]:	_____(jam)	_____(jam)	_____(jam)

a. Termasuk pasangan Anda jika Anda tidak memasukannya ke dalam anggota keluarga (e.g., suami/istri).

b. Masukkan siapapun dalam lingkungan kerja dalam kategori ini (e.g., Jika anda guru, masukkan siswa Anda sebagai teman kerja).

16a. Apakah Anda mencampur kata atau kalimat dalam bahasa yang berbeda ketika Anda berbicara? *(Termasuk, misalnya, memulai sebuah kalimat dalam satu bahasa tetapi menggunakan bahasa atau frasa dari bahasa lain di tengah-tengah kalimat.)* (Lingkari salah satu)

Ya/Tidak

16b. Jika Anda menjawab "Ya" pada pertanyaan 16a, maka sebutkan bahasa yang Anda campur dan perkirakan frekuensinya dalam percakapan normal dengan kelompok orang berikut. Silahkan perkirakan berdasarkan skala berikut (lingkari angka dalam tabel):

	Bahasa 1	Bahasa 2	Frekuensi
Anggota keluarga			1 2 3 4 5 6 7
Teman			1 2 3 4 5 6 7
Teman sekelas			1 2 3 4 5 6 7
Teman kerja			1 2 3 4 5 6 7

17. Dalam bahasa mana Anda berkomunikasi paling baik atau merasa paling nyaman dalam hal mendengarkan, berbicara, membaca, dan menulis dalam lingkungan berikut?

	Mendengarkan	Berbicara	Membaca	Menulis
Di rumah				
Dengan teman				
Di sekolah				
Di tempat kerja				

18. Seberapa sering Anda menggunakan tiap-tiap bahasa yang pernah Anda pelajari dalam aktivitas-aktivitas berikut? Silahkan lingkari angka dalam tabel berdasarkan skala berikut.

Tidak pernah	Jarang	Kadang-kadang	Rutin	Sering	Terbiasa	Selalu
1	2	3	4	5	6	7

Bahasa	Berfikir	Berbicara pada diri sendiri	Mengekspresikan emosi[a]	Bermimpi	Aritmatika[b]	Mengingat angka[c]
	12 34567	1234567	1234567	1234567	1234567	1 2 3 4 5 6 7
	1234567	1234567	1234567	1234567	1234567	1 2 3 4 5 6 7
	1234567	1234567	1234567	1234567	1234567	1 2 3 4 5 6 7
	1234567	1234567	1234567	1234567	1234567	1 2 3 4 5 6 7

a. Termasuk berteriak, mengumpat, menunjukkan perasaan, dll.
b. Termasuk menghitung, menghitung kembalian, dll.
c. Termasuk nomor telepon, nomor KTP, dll.

19. Berapa persen teman Anda yang berbicara dalam bahasa-bahasa yang pernah Anda pelajari? *(Total persentase harus berjumlah 100%.)*

Bahasa	Persentase
	%
	%
	%
	%

20a. Apakah Anda merasa bahwa diri Anda bicultural atau multicultural? *(Termasuk, misalnya, tumbuh dengan orangtua atau keluarga dari budaya yang berbeda atau tinggal di lingkungan dengan budaya yang berbeda untuk jangka waktu yang lama.)* (Lingkari salah satu)

Ya/Tidak

20b. Jika Anda menjawab "Ya" pada pertanyaan 20a, maka budaya/bahasa mana Anda merasa paling sesuai dengan diri Anda? Beri nilai hubungan Anda dalam kategori berikut untuk tiap-tiap bahasa/budaya. Silahkan lingkari angka dalam tabel sesuai dengan skala berikut.

Tidak ada	Sangat lemah	Lemah	Sedang	Kuat	Sangat kuat	Ekstrim
1	2	3	4	5	6	7

Budaya/Bahasa	Gaya hidup	Makanan	Musik	Kesenian	Kota	Tim olahraga
	1234567	1234567	1234567	1234567	1234567	1234567
	1234567	1234567	1234567	1234567	1234567	1234567
	1234567	1 2 3 4 5 6 7	1234567	1234567	1234567	1234567
	1234567	1234567	1234567	1234567	1234567	1234567

21. Silahkan berikan komentar Anda di bawah jika ada jawaban tambahan dari pertanyaan-pertanyaan di atas yang Anda rasa paling mendeskripsikan latar belakang atau penggunaan bahasa Anda?

22. Silahkan berikan komentar Anda di bawah untuk informasi lainnya mengenai latar belakang atau penggunaan bahasa Anda?

Appendix B. Samples of consent form

Permohonan Izin Orangtua untuk Keikutsertaan dalam Penelitian

Anak Anda dimohon untuk dapat berpartisipasi dalam penelitian yang dilaksanakan oleh Feisal Aziez dari University of Pannonia, Hungaria. Penelitian ini bertujuan untuk mendeskripsikan perkembangan Bahasa Inggris santri di Pondok Pesantren Muhammadiyah Al-Furqon, Singaparna, Tasikmalaya selama satu tahun. Penelitian ini merupakan bagian dari penyelesaian studi S3 peneliti di universitas tersebut.

Dalam penelitian ini, anak Anda akan diwawancara dan diminta untuk mengisi beberapa kuesioner serta menulis dan berbicara dalam bahasa Inggris. Ketika berbicara dalam bahasa Inggris, anak Anda akan direkam dengan perekam suara.

Penelitian ini akan memakan waktu sekitar 30 menit dalam seminggu. Partisipasi anak Anda dalam penelitian ini tidak akan mengganggu proses pembelajaran yang berlangsung di pondok dan juga tidak akan mempengaruhi nilai anak Anda dalam pelajaran bahasa Inggris. Peneliti juga berharap dengan menulis secara rutin, kemampuan menulis bahasa Inggris anak Anda akan meningkat.

Hasil dari penelitian ini akan dipublikasikan dalam konferensi atau jurnal internasional. Namun, kerahasiaan identitas anak Anda akan dilindungi. Partisipasi dalam penelitian ini bersifat sukarela. Anak Anda berhak memilih untuk tidak berpartisipasi atau mundur dalam proses berjalannya penelitian ini.

Jika Anda memiliki pertanyaan mengenai partisipasi anak Anda dalam penelitian ini, atau Anda ingin menerima laporan hasil saat penelitian telah selesai, silahkan hubungi:

Nama peneliti	: Feisal Aziez
Institusi	: Multilingualism Doctoral School, University of Pannonia, Hungaria
Email peneliti	: feiaziez@gmail.com
No. Tlp/Whatsapp	: +6281391109919 atau +36202875523
Supervisor	: Prof. Marjolijn Verspoor
Email supervisor	: m.h.verspoor@rug.nl

Pernyataan Pemberian Izin

Saya telah membaca surat permohonan izin dan saya telah memahami isi dari surat tersebut Saya secara sukarela memberikan izin agar anak Saya dapat berpartisipasi dalam penelitian ini. Saya juga memberikan izin untuk merekam anak saya dalam proses wawancara.

Siswa,

Nama: [redacted]

Orangtua/Wali,

Nama: [redacted]

Pernyataan Pemberian Izin

Saya telah membaca surat permohonan izin dan saya telah memahami isi dari surat tersebut Saya secara sukarela memberikan izin agar anak Saya dapat berpartisipasi dalam penelitian ini. Saya juga memberikan izin untuk merekam anak saya dalam proses wawancara.

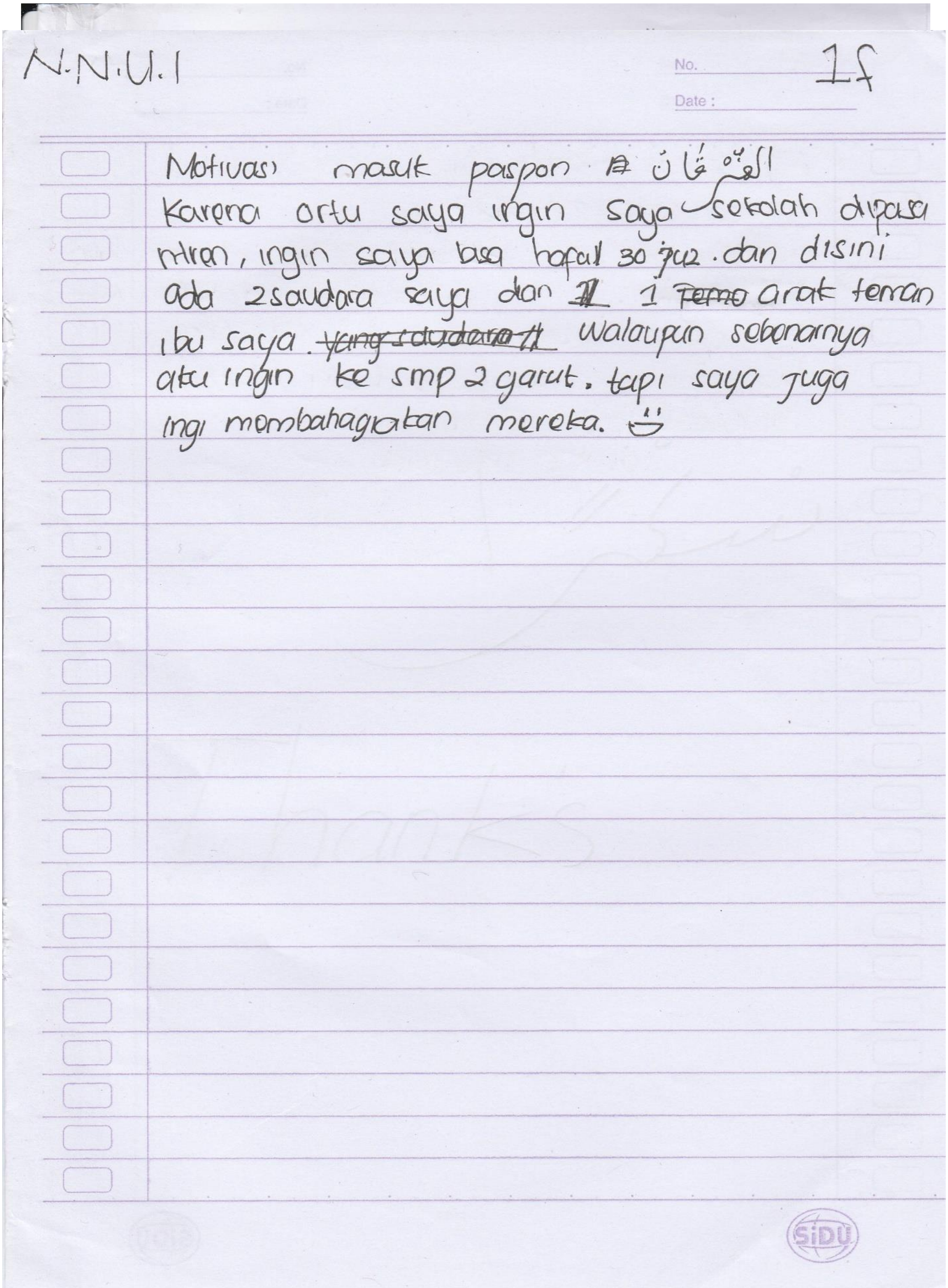
Siswa,

Nama: [redacted]

Orangtua/Wali,

Nama: [redacted]

Appendix C. Samples of learners' reflection on motivation



N.N.U.I

No. 1f

Date :

Motivasi masuk paspon الفرقان

Karena ortu saya ingin saya sekolah dipasa

ntren, ingin saya bisa hafal 30 juz. dan disini

ada 2 saudara saya dan 1 ~~Temo~~ anak teman

ibu saya. ~~yang saudara~~ walaupun sebenarnya

aku ingin ke smp 2 garut. tapi saya juga

ingi membahagiakan mereka. :)

G.A.N.

No.: Date:

Assalamu'alaikum

Saya Masuk Pesantren karna Suruhan Orang tua. Orang tua Saya Memasuk Pes Saya kepesantren karena Orang tua Saya ingin Membuat Saya lebih tau Ilmu agama islam dan bisa berbahasa arab dan inggris.

Awal nya saya gak mau masuk Pesantren. mau tidak mau saya harus masuk, dan akhirnya saya pun masuk Pesantren AL-FURQON.

OKEY

No.: C.M.R. Date: - 1F

Motifasi masuk al-Furqon

Saya masuk pesantren al-Furqon karena saya ingin menjadi anak ~~lagi~~ yang sholehah dan bisa membanggakan orang tua. Dan saya ingin lancar berbicara B. Inggris atau Pun B. Arab. Dan ingin mengadi anak yang berguna bagi agama islam. walaupun sebenarnya saya agak tidak ingin masuk pesantren tapi saya harus menuruti apa kata orang tua saya. Agar orang tua saya tahu bahwa saya dapat dipercaya oleh keluarga saya sendiri.

AMIN

KIKY Tommorow will be better

NAF

2F

No.

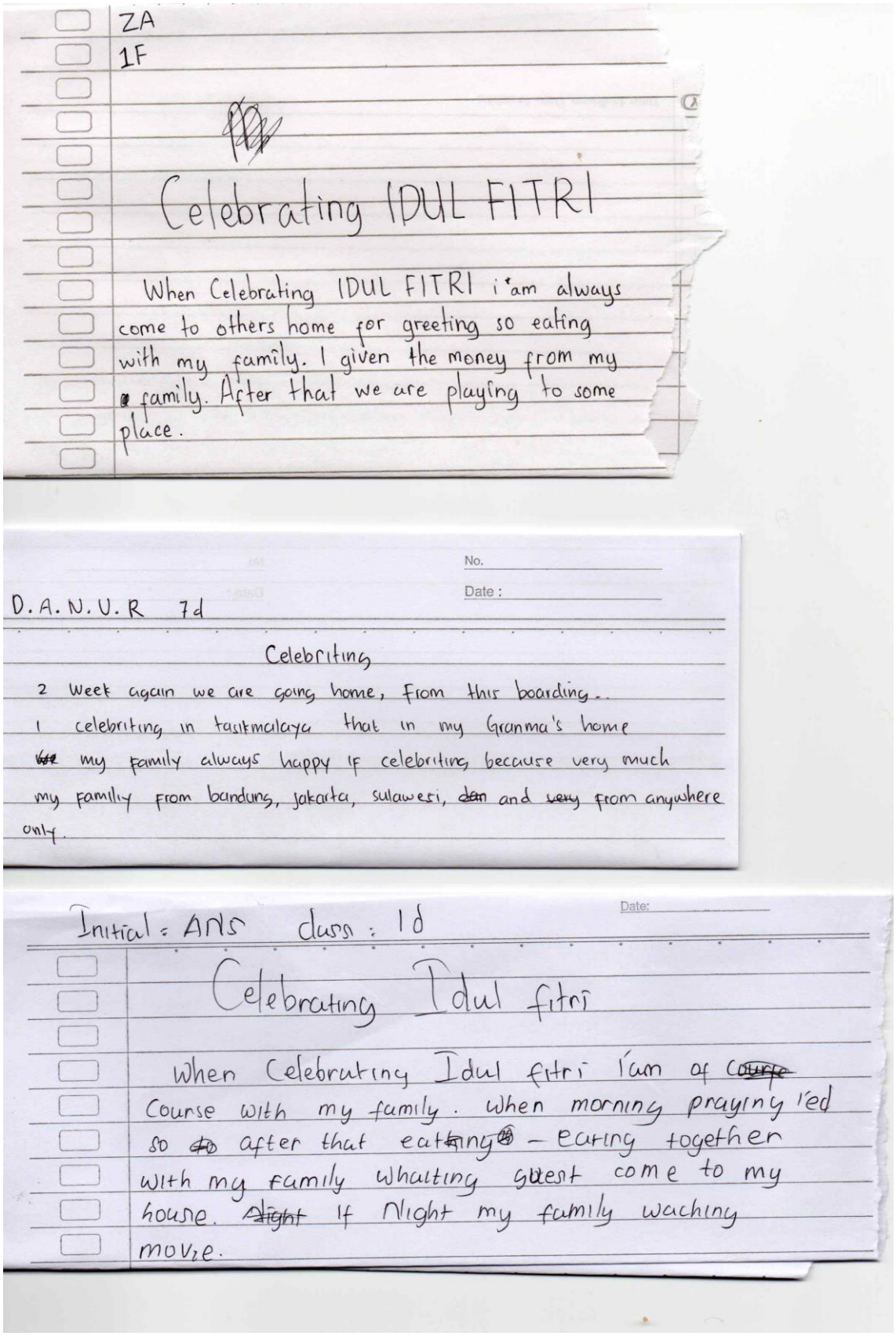
Date:

Motivasi saya masuk pondok karena saya melihat banyak orang yang pinter al-qur'an di lingkungan sekolah saya saya jadi ter - motivasi , dan karna saya juga termasuk orang yang suka malas jika menghafal sendirian , tapi jika bersama - sama saya lebih bersemangat , karna Itu dengan masuknya saya ke pondok saya jadi lebih bisa bersemangat belajar dan saya juga termasuk orang yang lemah dalam berbahasa karna Itu saya sangat ingin bisa bahasa jadi saya masuk ke pondok ini , dan yang paling penting saya ingin membanggakan orang tua , dan saya juga sangat suka kebersamaan, dan karna dirumah saya hanya punya satu kakak laki² jadi saat kesini saya merasa mempunyai banyak teman , sahabat dan keluarga yang selalu mendukung dan menemani setiap waktu.

^^

^o^

Appendix D. Writing samples



ZA
1F

Celebrating IDUL FITRI

When Celebrating IDUL FITRI i am always come to others home for greeting so eating with my family. I given the money from my family. After that we are playing to some place.

No.
Date :

D.A.N.U.R 7d

Celebriting

2 Week again we are going home, From this boarding..
I celebriting in tasikmalaya that in my Granma's home my family always happy if celebriting because very much my family from bandung, jakarta, sulawesi, dan and very from anywhere only.

Initial : ANS class : 1d
Date:

Celebrating Idul fitri

When Celebrating Idul fitri i'am of course Course with my family. when morning praying i'ed so after that eating - eating together with my family whaiting guest come to my house. night if night my family waching movie.

No.

Date :

NFMR
VIII D

Celebrating IDUL FITRI

every Idul Fitri . I go to my grand ma's home ~~at~~ and ~~getting get~~ meet with my familly.
but now, this year, I'm finished my Ramadhan moon in my boarding. cause I must do my examination. And I want do I'tikaf. so come for celebrating Idul fitri in H-3.

No. SF

Date : 8D

Celebrating idul Fitri

~~the~~ i go to my grand mather house over there i eating many foods, i like it :)
over there many body pupils and i like because very exiting. ~~ver~~
i play with my family until night because that time is very ~~s~~ different with nouther time

CELEBRATING IDUL FITRI

Idul Fitri will be coming soon !!! I'M Very waiting this Moment. Because In IDUL FITRI I get Much Money, and all my family Meet. And every IDUL FITRI Very Much foods. from Biscuit, cake, ketupat, etc. all pupils are asking apologize. to family, neigh bour, friend etc. Last Idul fitri ~~my family~~ i not Meet with all My family. Because they are Bussy. I hope this Idul fitri I Meet with all My family.

ZNY
Female
VIII D

N.R.M My Ramadhan's activity VII C

my Ramadhan's activity is so much. From 1st Ramadhan until 20th Ramadhan im in boarding. From 21st until the last in my home

Initial = R.N.D.P
Class = I F (C)

My Ramadhan's Activity

My Activity in Al - Furqor is L.E ngububurat buy Omlet buy ~~nodtle~~ ~~noodle~~ noodle, playing with my Friend and this is very amazing Fighting with uwef i'm everyday ~~Fing~~ ~~FTh~~ fighting with she is

RHR
7c Famale

MY RAMADHAN ACTIVITE

My ramadhan activite is sleeping, reading holy qur'an Play with my friend show Launge exibition, Pray to God, Sahur and break the past Pray tarawih, buy soteng in beloved cavetarion

initial : AZU
Class : VII C F

my Ramadhan's activity

In this Ramadhan in my boording much activity the first ramadh in my boarding ther is examination two week and after that ther is Languag Exebition (LE) 4 day . and if holiday I will sleep, wach etc. oh... yes forget I alwasy read holy book

4

My Ramadhan's Activity

My Ramadhan Activity are sleping, to help parent, story with friend, because we are in boarding school

---

DDPAZ
acf.

my Ramadhan's Activity.

my Ramadhan Activity is buy something that can make me satisfied. but buy something also not so much. usually my mind, so I make ramadhan Challans is khatam holly Qur'an. my Ramadhan's activity not just that is only but still much. maybe just that wich I can tell.

thanks :)

2AM
F

my Ramadhan's activity

I'm so much activity in this Ramadhan because I'm a more big more understand and must more good then before.

in this Ramadhan I'm so busy I'm so much a activity like a read holly book in the mosque so that I'm a khatam much then last years ago.

And my activity in this Ramadhan more much like queue up in the bathroom that make me tired if queue up .

and I'm a pray more, more, and more, and read holly Qur'an more, more, and more and my activity in this Ramadhan much again and etc

Appendix E. Excerpts of the vocabulary book

Your **prayers** are your **light**,
Your **devotion** is your **strength**,
Sleep is the **enemy** of both.
Your life is the **opportunity** that life can give you.
If you **ignore** it, if you **waste** it, you will only turn to **dust**.

1.	Dining room	Ruang Makan	مَطْعَمٌ
	I go to dining room to take my breakfast		
	أَذْهَبُ إِلَى المَطْعَمِ لِتَنَاوُلِ الإِفْطَارِ		
2.	Kitchen	Dapur	مَطْبَخٌ
	My mother is cooking rice in the kitchen		
	تَطْبَخُ أُمِّي رُزًّا فِي المَطْبَخِ		
3.	Plate	Piring	صَحْنٌ
	I wash my plate after having my breakfast		
	أَخَذَ مُحَمَّدٌ صَحْنَهُ قَبْلَ اِسْتِعْمَالِهِ لِلأَكْلِ		
4.	Spoon	Sendok	مِلْعَقَةٌ
	I buy the new spoon at Students Cooperation Section		
	غَسَلَ أَخِي الْمِلْعَقَةَ بَعْدَ الْأَكْلِ		
5.	Knife	Pisau	سِكِّيْنٌ
	I cut the meat with knife		
	حَمَلَ أَبِي السِّكِّيْنَ لِقَطْعِ لَحْمِ الْبَقَرَةِ		
6.	Fork	Garpu	شَوْكَةٌ

	I eat the vegetables with the fork		
	أَكَلْتُ الْمَكْرُوْنَةَ بِالشَّوْكَةِ		
7.	Food	Makanan	طَعَامٌ
	I buy some foods at Students Cooperation Section		
	أَشْتَرِى طَعَاماً فِي الجَمْعِيَّةِ التَعَاوُنِيَّةِ لِلطَّلَبَةِ		
8.	Drink	Minuman	شَرَابٌ/ مَشْرُوْبٌ
	I take some drink after breakfast		
	أَتَنَاوَلُ شَرَاباً بَعْدَ الإِفْطَارِ		
9.	Breakfast	Makan pagi	فُطُوْرٌ
	Did you have your breakfast?		
	أَ تَنَاوَلْتَ فَطُوْرَكَ؟		
10.	Lunch	Makan siang	غَذَاءٌ
	أَتَنَاوَلُ غَذَائِيْ بَعْدَ صَلاَةِ الظُهْرِ		
	I take my lunch after Dhuhur prayer		
11.	Dinner	Makan malam	عَشَاءٌ
	العَشَاءُ فِي السَاعَةِ السَابِعَةِ		
	The dinner is at seven o'clock		
12.	Refrigerator	Lemari es	ثَلاَّجَةٌ

	I freeze the ice in the refrigerator		
	أَضَعُ الثَّلْجَ فِي الثَّلاَّجَةِ لِيَبْرُدَ		
13.	Bathroom	Kamar mandi	مَّامٌ
	I clean the bathroom once a week		
	أَقْرَأُ الدُّعَاءَ قَبْلَ دُخُوْلِ الْحَمَّامِ		
14.	Water scoop	Gayung	رَفَةٌ
	Pass the water scoop to me please!		
	اشْتَرَى فَرِيْدٌ الْمِغْرَفَةَ لِلاِسْتِحْمَامِ		
15.	Towel	Handuk	شَفَةٌ
	Hang my towel after using please!		
	أُجَفِّفُ الْمِنْشَفَةَ بَعْدَ الاسْتِعْمَالِ		
16.	Tooth brush	Sikat gigi	جُوْنٌ
	I clean my teeth before sleeping with a tooth brush		
	أُنَظِّفُ أَسْنَانِيْ بِالْفِرْجَوْنِ قَبْلَ كُلِّ صَلاَةٍ		
17.	Tap	Kran	فِيَّةٌ
	The water flows from the tap		
	يَخْرُجُ الْمَاءُ مِنَ الحَنَفِيَّةِ		
18.	Tooth paste	Pasta gigi	جُوْنٌ

Buy the tooth paste for me!			
هَلْ مُمْكِنُ أَنْ أَطْلُبَ مَعْجُوْنَكَ؟			
19.	Soap	Sabun	صَابُوْنٌ
Use the soap to make your body fresh			
مِنْ فَضْلِكَ! أَنْ تَأْخُذَ لِي الصَّابُوْنَ فِي الْخِزَانَةِ			
20.	Student's Cooperation	Koperasi Pelajar	جَمْعِيَّةٌ تَعَاوُنِيَّةٌ لِلطَّلَبَةِ
I visit the student's cooperation for buying my necessary			
اشْتَرَيْتُ الأَدَوَاتِ الْمَدْرَسِيَّةِ فِي الْجَمْعِيَّةِ التَّعَاوُنِيَّةِ لِلطَّلَبَةِ			
21.	Library	Perpustakaan	مَكْتَبَةٌ
I want to look for the Islamic book in the library			
أَزُوْرُ الْمَكْتَبَةَ لِبَحْثِ عَنِ الْكُتُبِ الإِسْلاَمِيَّةِ			
22.	Dormitory	Asrama	مَسْكَنٌ
The students are staying in the dormitory			
يَسْكُنُ الطَلَبَةُ فِي المَسْكَنِ (المَسَاكِنِ)			
23.	Mattress	Kasur	فِرَاشٌ
I have the thick mattress for sleeping			
أَسْتَعْمِلُ الْفِرَاشَ الثَّخِيْنَ عِنْدَ النَّوْمِ			
24.	Pillow	Bantal	وِسَادَةٌ

I use the pillow to sleep well			
رَتِّبِ الْوِسَادَةَ عَلَى الْفِرَاشِ!			
25.	Blanket	Selimut	بَطَانِيَّةٌ
The blanket is useful in the cold weather			
اِسْتَعْمِلِ البَطَانِيَّةَ لِدَفْعِ الْبُرُوْدَةِ فِي اللَّيْلِ!			
26.	Cupboard	Almari	خِزَانَةٌ
I paint my cupboard by the blue color			
أَصْبَغُ جَمِيْعَ الْأَدَوَاتِ الْيَوْمِيَّةِ فِي الْخِزَانَةِ			
27.	Hanger	Gantungan baju	مِشْجَبٌ
I hang my clothes with hangers			
أُعَلِّقُ لِبَاسِيْ بِالْمِشْجَبِ			
28.	Cloth	Kemeja	قَمِيْصٌ
I want to iron my cloth			
أَكْوِى الْقَمِيْصَ لِيَكُوْنَ مُرَتَّباً			
29.	Trousers	Celana	سِرْوَالٌ
That trouser is too long for me			
هذَا السِّرْوَالُ طَوِيْلٌ عَلَيَّ			
30.	Sarong	Sarung	إِزَارٌ

Moslems are wearing the sarong when prayer			
أَسْتَعْمِلُ الْإِزَارَ عِنْدَ كُلِّ صَلَاةٍ			
31.	Belt	Sabuk/ Gesper	حِزَامٌ
Fasten your belt to make your sarong tighter			
أَرْبِطُ الْإِزَارَ بِالْحِزَامِ عِنْدَ الصَّلَاةِ			
32.	Shoes	Sepatu	حِذَاءٌ
I blacken my black shoes by the shoes polish			
أَضَعُ الْحِذَاءَ فِيْ رَفِّ الْحِذَاءِ بَعْدَ الْإِسْتِعْمَالِ			
33.	Socks	Kaos kaki	جَوْرَبٌ
I wash my socks once a week			
أَغْسِلُ الْجَوْرَبَ مَرَّةً فِي الْأُسْبُوْعِ			
34.	Money order	Wesel	حَوَالَةٌ
My parent sends me money order monthly			
أَرْسَلَ لِيْ أَبِيْ الْحَوَالَةَ شَهْرِيًّا			
35.	Parcel	Paket	رِزْمَةٌ
Ahmad gets the parcel from his mother			
إِسْتَلَمَ أَحْمَدُ الرِّزْمَةَ مِنَ الْبَيْتِ			
36.	Letter	Surat	رِسَالَةٌ

I receive the letter from my mother			
تَسَلَّمْتُ الرِّسَالَةَ مِنْ وَالِدَتِيْ			
37.	Letter box	Kotak Surat/ Kotak Pos	صُنْدُوْقُ الْبَرِيْدِ
You put your letter into the letter box			
أُدْخِلُ الرِّسَالَةَ إِلَى صُنْدُوْقِ الْبَرِيْدِ بَعْدَ كِتَابَةِ الْعُنْوَانِ			
38.	Postman	Tukang pos	سَاعِي الْبَرِيْدِ
The postman sends the letters to the houses			
أَرْسَلَ سَاعِي الْبَرِيْدِ الرِّسَالَةَ إِلَى الْبُيُوْتِ			
39.	Field	Lapangan	مَلْعَبٌ
The students watch the competition in the field			
لَعِبَ الْأَوْلَادُ كُرَةَ الْقَدَمِ فِي الْمَلْعَبِ			
40.	Football/ soccer	Sepak bola	كُرَةُ الْقَدَمِ
I watch football match			
أُشَاهِدُ مُبَارَاةَ كُرَةِ الْقَدَمِ			
41.	Volleyball	Bola voly	كُرَةُ الطَّائِرَةِ
Volley ball is my hobby			
لَعِبَ الْأُسْتَاذُ كُرَةَ الطَّائِرَةِ فِي الْمَيْدَانِ			
42.	Table tennis/ ping-pong	Tenis meja	كُرَةُ الطَّاوِلَةِ