Chapter 4

Input outside the classroom and vocabulary development: A dynamic perspective

4.1 Introduction

Both the frequency (quantity) and quality of the input are crucial for the successful development of a second language. However, in many classrooms around the world, English language instruction often occurs within a non-authentic context. In Taiwan, language learning syllabi are constructed around the grammar-translation method, and the main focus is on word-list translation or textbook reading. Whilst students attend eight hours of English classes a week, they rarely hear a complete English sentence or produce any spoken English themselves. Most of the in-class English input is in fact presented in Chinese. This level of in-class English input is not quantitatively sufficient in terms of the number of repetitions necessary to learn new vocabulary. In addition, it does not qualitatively support contextual comprehension of the words.

Receiving input outside the classroom could potentially overcome the shortcomings of such English classes in Taiwan, but most of the Taiwanese English learners are not aware of the available access to input outside the classroom (e.g. smart phone or computer). Exposure to English in an unstructured setting increases the amount of input learners receive, which in turn increases the number of words learners are exposed to. Learners can watch English episodes, listen to English radio, read online news, play online games, and chat on Facebook. These sources combined with the ubiquity of TV, computers, ipads, or smart phones have resulted in an environment where learners can receive English input whenever and wherever they choose. These rich and varied opportunities could potentially lead to improvement of English proficiency beyond the learners’ in-class learning.

The importance of input outside the classroom has been emphasized in several studies. Xu (2010) found that input outside the classroom is influential in supporting the retention of English words. Dutch learners, who received lots of English input in everyday life, had a higher retention rate of English words than Chinese learners, who had limited access to authentic English input in everyday life. Another study in the Netherlands (De Bot, De Quay-Peeters, & Evers, 2004) elucidated the factors that determine the English proficiency of thirteen to sixteen-year old lower secondary pupils. 1574 pupils were tested on four skills of English proficiency: listening ability, reading ability, speaking ability, and writing ability. This study showed that the prime determiner of English proficiency was not school-based teaching but rather exposure to English through the media (TV, music, computers, and the internet) outside of the classroom. Listening to pop songs, especially when the listener paid attention to the lyrics, had a strong effect on their aural comprehension skills. Watching TV programs, with either English subtitles or spoken English, also had an impact on language proficiency, especially on listening ability.

These findings reflect the main argument of Ellis’ (2002) and Larsen-Freeman’s (2002) studies: input, operationalized as the frequency of occurrence, is key in successful language acquisition. The ability to use English fluently can be seen
as a cumulative effect of interaction within an authentic English environment. The more learners connect themselves with a naturalistic English context (environment), the more likely they are to develop a good command of English. It is not likely that a learner will be able to learn a language in isolation without authentic English interaction. With increased amounts of contact time with English through various channels outside the classroom, learners enlarge the possibility of successfully learning English.

However, the large-scale quantitative studies reported earlier have not been able to provide a satisfactory explanation of what exactly occurred during the language learning process. The explanatory power of the study by De Bot, De Quay-Peeters, and Evers (2004) was quite limited as only 14% of the variance was explained. Moreover, previous studies have concentrated on the outcomes of the process (i.e. the product), rather than on the process itself. Therefore, a more detailed study is needed, in which learners are followed intensively and longitudinally to observe the development of language over time. It is not feasible to carry out such research at that large scale due to limited time and resources. A smaller scale longitudinal study should be able to give us sufficient insight into the real-time process of language learning as it is influenced by input outside the classroom. The developmental process can be investigated by the longitudinal data with the application of a dynamic mathematical model, the logistic model, which can be empirically and theoretically considered an ideal model to describe the vocabulary development.

Our study intends to answer three research questions:
1. Do learners of English improve their writings (operationalized as lexical difficulty) when they receive more exposure from input outside the classroom?
2. Do English learners’ writings deteriorate (operationalized as lexical difficulty) when they receive less exposure from input outside the classroom?
3. Can the logistic model meaningfully interpret the development of lexical difficulty for each individual learner?

We followed four Taiwanese learners of English longitudinally for five months, which produced a time series of fifty-six writings. These writings were analysed with respect to the learners’ development of lexical difficulty. We then compared the writing development (in terms of lexical difficulty) with the differing degrees of amount of input the learners received outside the classroom. This paper begins with a general review of research into vocabulary learning and focuses on the differences between our study and previous studies on this issue, followed by a focus on vocabulary learning from a dynamic systems theory perspective. The description of data collection combines detailed information regarding the participants, the materials, and the procedures with the time series of writing production. Finally, the analyses and results pertaining to the concepts of dynamic system theory lead to the discussion and conclusions about the application of a dynamic approach to investigating the influence of input outside the classroom on lexical difficulty.
4.2 Theoretical background

4.2.1 Vocabulary Size and Lexical difficulty

Vocabulary has long played an essential role in studies of second language acquisition. Learners of English need around 8000 to 9000 word families (all morphologically related lexical items are counted as one family) or about 98% coverage of the vocabulary, to read authentic texts such as novels or newspapers (Nation, 2006). As for listening comprehension, they need about 90% coverage of the vocabulary to sufficiently understand the spoken discourse (Laufer, 1989). A large number of lexical items are required to read and comprehend English as a second language. The number of lexical items, referred to as vocabulary size, is regarded as an important indicator of the capacity to operate in English.

The Lexical Frequency Profile (LFP), developed by Laufer and Nation in 1995, is an approach to estimating the productive vocabulary size of English learners from their writings. LFP generates a profile of the writing productions by describing the lexical content of the text in terms of frequency bands. For example, a profile of a writing production could report that 92% of the vocabulary falls in the frequency band 1 to 1000 (1k), 5% of the vocabulary falls in the frequency band 1001 to 2000 (2k), and 3% of the vocabulary falls in the frequency band 2001 to 3000 (3k). Calculating the percentage of the vocabulary that falls into each frequency band makes it possible to characterize each text in a standardized way.

Meara (2005) used the percentage of words in 1k, 2k, and 3k (2001+) based on the study of Laufer and Nation (1995) to generate the simulated data to test the sensitivity of LFP in different ranges of vocabulary size. It is suggested that 100% of the possibility that the LFP can distinguish the estimation of vocabulary size between 2000 and 3000, that 90% of the possibility that the LFP can distinguish the estimation of vocabulary size between 2500 and 3000, and that 67% of the possibility that the LFP can distinguish the estimation of vocabulary size between 3000 and 3500. As our participants’ vocabulary inferred from the beginning level of GEPT test was at least 2800 words in size but not exceeding 4000 words as none of them passed the intermediate level of GEPT test. So, for this range (2800-4000), it seems that LFP cannot well distinguish the writings produced by our participants. Similar to the Meara’s (2005) finding, Edwards and Collins (2011) confirmed the fact that the use of 1k word proportions in LFP to distinguish learners of different vocabulary sizes becomes less effective as vocabulary size increases. The change of probability of 1k words diminishes when vocabulary size increases. Using LFP as a tool to assess our participants’ vocabulary size may not be suitable.

$V_{\text{size}}$ (Meara & Miralpeix, 2004) is an alternative approach to estimating productive vocabulary size from writing samples. The assumption of $V_{\text{size}}$ follows the assumption of the Zipf’s law (1935, 1949), where words are acquired in order of their frequency, and the frequency of the occurrence of the words is roughly
 inversely related to the rank of the word. A more proficient learner would use more low frequent words; a less proficient learner would use fewer low frequent words. The following equation describes such property of vocabulary learning.

\[ F = \frac{KT}{R} \]  

(4.1)

where \( F \) is the frequency of a word; \( R \) is its rank; \( T \) is the length of a text; \( K \) is a constant of proportionality. The equation is transformed into the logarithmic weighting (equation 2) since words are not likely to occur equally in a text. This transformation makes the relationship between \( \log(R) \) and \( \log(F) \) a linear one. A detailed description of how Zip’s law is used to estimate the vocabulary size can be found in Edward and Collins’ (2011).

\[ \log(F) = \log(K) + \log(T) - \log(R) \]  

(4.2)

After loading the written text into V_size, we first get a word list and the percentage of the words in each frequency band (Frequency band A, from 1 to 500; Frequency band B, from 501 to 1000; Frequency band C, from 1001 to 1500; Frequency band D, from 1501 to 2000; and Frequency band E, above 2000). Then the V_size program finds the best matching pattern of the data based on Zipf’s law and generates the estimated value of the vocabulary size of this writer.

However, there are some problems of using V_size. Firstly, the percentage of the words in each frequency band depends largely on the dictionary used. If we wish to measure the V_size of writings from the same learner, we must use the same dictionary as a reference to the percentage distributed over the frequency band. Second, as our participants are beginner learners of English, they tend to make some errors in their writing productions. The treatment of these errors should be consistent and should eliminate the effect of the error on the estimation of V_size.

There are also restrictions of using V_size. As V_size is sensitive to the words of low frequency band (E), V_size seems to behave erratically. To have a rough estimation of the sensitivity, we took one sample written text with 131 words to observe how V_size may fluctuate. A sample written text of 131 words with 79% words in A (1-500), 10% words in B (501-1000), 7% in C (1001-1500), 4% words in D (1501-2000), and 0% words in E (above 2000) is estimated to be written by a learner whose V_size is 2100. An increase from 0% to 1% in E (2 more words in E) enlarges the V_size by 200 (2100 to 2300). An increase of 2% (3 or 4 words in E) enlarges the V_size by 300 (2100 to 2400). An increase of 3% (5 words in E) enlarges the V_size by 500 (2100-2600). An increase of 4% (6 words in E) enlarges the V_size by 700 (2100-2800). An increase of 5% (7 words in E) enlarges the V-size by 1000 (2100-3100). This means that even with a small number of words
used in low frequency band E, the value of $V_{\text{size}}$ changes erratically. If the text in itself is very short, the small change of the number of words will largely affect the percentage $V_{\text{size}}$ and thus the $V_{\text{size}}$ will be of high variability. Moreover, if the topic requires the use of many infrequent words, $V_{\text{size}}$ will also be of high variability.

Due to the high variability of $V_{\text{size}}$, we would take the estimated vocabulary size generated by $V_{\text{size}}$ as an index of lexical difficulty in writing. Although lexical difficulty has been found not only related to the frequency band but also related to the difficulty of forming and image or the length of the word (Dunn, 2013), $V_{\text{size}}$, if used as an index of lexical difficulty, can at least reveal the lexical difficulty in terms of frequency band. The higher the value of $V_{\text{size}}$ is, the higher the lexical difficulty is. As we assume that the effect of the exposure to input outside the classroom may bring about immediate effects on the use of more difficult (low frequency) words, we will use $V_{\text{size}}$ as an index of lexical difficulty to observe this development over time.

4.2.2 Intentional and incidental vocabulary learning

To enlarge the number of the lexical items, two approaches to learning vocabulary have been largely investigated: intentional and incidental vocabulary learning. Intentional vocabulary learning (explicitly drawing the learners’ attention to the lexical items themselves) is especially effective for the retention of words. Smith (2004) found that learners retained more novel lexical items that were learnt in an Internet chat program, in which words were largely used through the interaction between the subjects. Walter & Bozkurt (2009) also found that learners have better retention of words when the words on the notebook are firstly previewed and later incorporated in the classroom. Maximizing the engagement with the words is the main principle in designing intentional learning activities (Schmitt, 2008).

On the other hand, incidental learning (implicitly exposing learners to meaning-focused input) increases the number of word repetitions and allows learners to consolidate their contextual knowledge of the vocabulary. Hulstijn and Laufer (2001) investigated the incidental learning that results from reading. They compared the short-term and long-term receptive vocabulary retention between three levels of involvement loads, reading comprehension, comprehension plus filling in target words, and composition-writing with target words. Their results showed that words which were processed with composition-writing were more successfully retained than with comprehension or comprehension plus filling in target words. Knight (1994) also found that Spanish intermediate-level language learners increased their English vocabularies by reading and deducing meaning through guessing or using a dictionary. In light of these studies, it can be argued that acquisition of vocabulary could be enhanced by intentionally or incidentally engaging learners with the words to greater degrees.
These studies, focusing on either intentional or incidental learning, mostly investigate what type of “in-class input” learners encounter and draw conclusions based on how many lexical items learners are able to recognize or produce after the treatment. However, the potential impact of natural exposure to English input on promoting the acquisition of vocabulary, especially with in-depth longitudinal observations, remains largely unknown. Input outside the classroom may potentially enhance vocabulary learning by exposing learners to the words more frequently, which would maximize their lexical engagement in a context-based environment and raise the chances for learners to retain and then produce the words themselves. Therefore, there is a need for a study that investigates the potential effect of input outside the classroom on vocabulary acquisition.

4.2.3 A dynamic perspective of vocabulary development

The application of dynamic system theory (DST) to the theory of second language development includes several basic characteristics (De Bot & Larsen-Freeman, 2011). One of the starting points of a dynamic system approach is that there is a sensitive dependence on the initial conditions of the individual. It features a non-linear development with interconnected variables, and these variables change through internal reorganization and interaction with the environment. “Dynamic”, in fact, refers to the changes that occur in a system as a result of internal forces and “energy from outside itself” (De Bot & Larsen-Freeman, 2011, p.8).

DST has been applied to first and second language learning (Cameron & Larsen-Freeman, 2008; Van Geert & Steenbeek, 2005) and human cognitive systems (Elman, 2004; Spivey, 2007). These studies intend to depict the tendencies, patterns, and the contingencies of language development, which are iterative, non-linear, time-dependent, and interconnected: the key characteristics of DST.

4.2.4 Modelling vocabulary development

One way of investigating dynamic development is to build a mathematical model that quantifies the dynamic relationships between the variables. To apply a mathematical model, we start by observing the developmental patterns of the data. This helps us determine which mathematical model we could choose to describe the trajectory. The trajectory generated from the model must resemble the trajectory of the data. Second, we look into the theoretical background of the model, which must be compatible with the concepts on which our data is based: iterativity, non-linearity, and time-dependency.

Vocabulary development as a dynamic non-linear process, is iterative, and is limited by the availability of resources. The size of the vocabulary in time “\( t + 1 \)” depends on time “\( t \)”: each data point depends on the previous data point. The
way in which the variables interact unfolds step by step. Learners have a limited learning capacity (working memory), limited input (external resource), or limited motivation; thus the vocabulary development does not increase infinitely but stagnates in the end. The logistic model presents a dynamic theoretical equation where it models iterative growth and allows us to model on the basis of initial conditions, carrying capacity, and learning rate.

The logistic model was originally used to describe the population growth over time by Verhulst (1845). The population $P(t)$ varying over time can be described as:

$$\frac{dP(t)}{dt} = rP(t)(1 - P(t)/K) \quad (4.3)$$

The left part of the equation demonstrates that the population ($P$) changes with the time ($t$). The right part of the equation demonstrates that the growth rate of the population ($r$) changes with the time and that the population ($P$) is limited by its resources ($K$).

The pattern of the logistic model is likely to match that of vocabulary development: the development is slow in the initial state, is faster in the middle state, and is slow again when the maximum capacity is approached in the final state. This is illustrated in Figure 4.1. The growth pattern changes with three variables: the initial value (the beginning value of the development), the learning rate (the slope of the development), and the carrying capacity (the maximum final value of the development). We illustrated how the adjusted values of three variables change the trajectory in Figure 4.1. When changing the value of the carrying capacity ($K$) from 1 to 1.5 and holding other parameters constant, we find the difference between line 3 and line 1 (compared to line 3 as a standard). Changing the value of the learning rate from 0.3 to 0.4 and holding other parameters constant, we find the difference between line 3 and line 4. Changing the initial value (the beginning value of the development) from 1.2 to 2.0 and holding other parameters constant, we find the difference between line 3 and line 2.

Previous applications of logistic equation modelling have shed important light on the dynamic interactions of components within dynamic systems. Caspi (2010) utilized the logistic model to describe the academic vocabulary development of four advanced learners of English. She found that the more receptive vocabulary knowledge acted as a conditional precursor for the more productive vocabulary knowledge. Chan (2010) used the logistic model to describe the vocabulary development of three different levels of vocabulary knowledge for learners of English. In both studies, the logistic model successfully interprets the vocabulary development of different levels of vocabulary knowledge using three variables: initial value, learning rate, and carrying capacity. The carrying capacity represents all of the learner’s resources in the time frame.
4.3 The study

4.3.1 Method

We followed four beginning learners of English for five months through their log of input outside the classroom and their fifty-six writings. Based on their logs and the input given by the experimenter, we defined low, intermediate, and high degrees of input outside the classroom, operationalized as hours of English input per week. The learners freely produced two to three writings per week without any time limit on a Facebook writing club that was set up for the experiment. Measuring the lexical difficulty (with the measure of \(V_{\text{size}}\)) from their writings, we observed the general trajectory of the data, smoothed it, fitted the smoothed data with the logistic model, and then obtained the value of the three variables. We also compared the mean \(V_{\text{size}}\) with high, intermediate, and low degrees of input exposure.

4.3.2 Participants

Four Taiwanese beginning learners of English participated in the study: Skid, Grace, Gloria, and Tina. At the start of the experiment, their receptive vocabulary size was approximately 2260 words, according to the General English Proficiency Test they passed (Wu, 2012). This figure (2260) provided us with a rough idea of the English proficiency level of these learners. Through interviewing the subjects’
English teacher, several details were already known about the participants. Skid (age 32) could read well but could not comprehend fast-paced English conversation, speak fluently, or write smoothly. Tina, Grace, and Gloria (age 15) were very close friends. They had an equal capacity in four aspects of English: speaking, listening, writing, and reading. Tina was particularly good at speaking fluently, and Gloria was particularly good at writing error-free sentences.

4.3.3 Materials

There are five material resources the four learners made use of. They read online news written for the ESL learners on “Breaking News English”, watched free online movies in English with Chinese subtitles on “Tube+”, watched the latest English music videos on “YouTube”, chatted in English with friends on “Facebook”, and listened to English songs with “Itunes.” Among the five resources, movies and readings were the two main types of English input.

In this study, the larger part of the input was experimentally manipulated; some of the input was selected by the participants themselves. We kept track of the learners’ input from their logs, which convey the history of input outside of the classroom. They wrote down what type of input (movie), how long they were exposed to the input (2 hours), and what they did with the input (e.g. I looked up two words from the movie). The fifty-six written language samples were analysed with respect to the learners’ development of lexical difficulty.

4.3.4 Input stages

From the participants’ logs, we defined three degrees of input outside the classroom: low (less than 2 hours of English input per week), intermediate (2 to 5 hours of English input per week), and high (more than 9 hours of English input per week). In five months, Skid went from a low level input to an intermediate level of input, mostly receiving the input from movies and reading. Tina, Grace, and Gloria spent most of their time together and went from a low level of input to an intermediate level (mostly through reading), then a high level, and then back to an intermediate level (mostly through movies).

In the first stage, classified as low input, where Skid completed writing 1 to 19, he did not have any English input. In the second (intermediate) stage, from writing 20 to 38, he mostly watched movies as English input. In the third (still intermediate) stage, from writing 39 to 56, he mostly read online news as English input.

In their first low-input stage, from writing 1 to 10, Tina, Grace, and Gloria only had English songs as their English input. In the second (intermediate) stage, from writing 11 to 20, they mostly read online news as their main English input. In the
third stage, from writing 21 to 37, they read online articles everyday, watched two to three English movies a week, and frequently chatted with friends on Facebook or in person in English. This was their stage with a high-level of input. In the fourth stage, where they fell back to an intermediate level of input, writing 38 to 56, they watched two English movies per week as their main source of English input.

4.3.5 Writing tasks

The four participants wrote approximately 200 words for each writing. The topics of the 56 writings are all TOEFL\(^1\) writing topics. These topics are related to expressing opinions on agreeing or disagreeing with a statement, or about comparing the advantages and disadvantages of a subject. For example, they produced one writing piece on comparing the advantages of having friends who are different from them and the advantages of having friends who are similar to them. They also produced another type of writing piece on whether they agree parents are the best teachers. For such intensive longitudinal study, it is important to maintain the difficulty of different writing topics at the same level. These topics, highly related to the learners’ life experience, could potentially elicit equally as many words from each learner. Although the topics varied, they were invariably general in nature and required little topic-specific vocabulary. Moreover, as our participants have a comparatively small vocabulary pool, the topics were limited in word choice: many of the words used largely overlapped.

4.3.6 Procedure

Over five months, the four participants posted two to three 200-word writings per week on a Facebook writing club constructed for the purpose of the experiment. Participants were able to read each other’s posts and to comment on what they read or wrote. As could be observed from the content of the writings, the four participants demonstrated very little imitation of each other’s work. Some of the participants may have learnt new vocabulary from the productions of the other participants, but this new vocabulary could also have been learnt from alternative English input. Regardless, there was no significant evidence of imitation apparent in the writings. The participants could choose whether or not to respond to each other if they received feedback on their posts. There was no requirement to correct their grammatical or lexical errors.

All writing productions were transformed into text files, so we could utilize V\_size\(^1\) v2.0 (Meara & Miralpeix, 2004) to estimate the lexical difficulty with the value

\(^1\)TOEFL writing topics can be found in the following link: http://www.ets.org/Media/Tests/TOEFL/pdf/989563wt.pdf
generated from this tool: vocabulary size. In our study, the participants made the following types of errors: subject-verb agreement, plural noun, misuse of categories of words (e.g. adjective and adverb), parallelism (balance and equality especially in tenses), verb form and tense, orthography, sentence structure, misuse of word meaning, preposition, pronoun, and articles. Among these errors, misuse of categories of words, verb form, orthography, and misuse of word meaning may influence the $V_{size}$ to a greater degree as they are related to the word count in the frequency band. We count in the misuse of categories of words and verb form as at least the partially productive word form and meaning are achieved and as this type of error is related to their typical L1-L2 translation learning style. We did not count in the misuse of word meaning as only the productive word form is learned. As for mis-spelling, if the error was consistent, we did not count it in. If the error was not consistent, implying that it could be the typing problem instead of orthography, we count it in. Proper names, like Taiwan and Japan, were re-coded as the highest frequency band to avoid an overestimation of the productive vocabulary size; hyphenated words, (e.g. “well-known”), are counted as two words in $V_{size}$, and thus we edited these down to one word.

Here is one example how we clean up the data:

I think that a nice bridge is really important for a country! Because it may bring a lot of convenient (error of incorrect use in the sentence), people can go anywhere easily, even in the big country! I hate traffic jam so much. That is too tiring and boring, and it also waste (error of incorrect use in the sentence) our time. That is not a good situation, I think! I think the most terrible traffic jam I have seen was in Seoul (proper name)! Ten lines were full of cars! I don’t (abbreviation) know that I have to say it’s (abbreviation) beautiful or horrible! But I think that the people in the car were so poor! What time will be arrived (error of incorrect use in the sentence)? That is a good question! Fortunately, we have MRT (proper name) in Taipei, it brings a lot of convenient! I think without MRT, Taipei’s (proper name) traffic jam will be better than Seoul (proper name)! And hopefully, there wasn’t (abbreviation) any big problem about traffic jam in Taiwan, except the Festival day (error of incorrect use in the sentence). But Taiwan is too small. I think this is not a big problem for us! And we also have a lot of transportation vehicle (error of incorrect use in the sentence), so I just want to say, we are lucky to live in Taiwan!

4.3.7 Design and analyses

We assume that our four participants would increase their lexical difficulty ($V_{size}$) whilst they received more input outside the classroom and that the development of lexical difficulty would be stagnant when there was no exposure to input outside the classroom. We closely observed the learners’ amount of input outside the classroom from their logs and analysed their writings in terms of lexical difficulty. In order to construct a dynamic concept of the development, each individual, taken
as a constantly changing complex system, has a set of variables that interact over time. In our study, two variables are selected for observation: the amount of input outside the classroom, and the written language samples, operationalised as lexical difficulty. The basic assumption is that with more contact with an English context, learners may increase their lexical difficulty over time, while with less input outside the classroom, learners’ lexical difficulty will be stagnant or even decrease. Input outside the classroom interacts with the writing productions over time, causing change in the individual’s complex system. Input outside the classroom promotes the use of more advanced words in the writings, and in turn the writings, consolidating the use of the words, facilitate the comprehension of the input outside the classroom in the next time point. The previous output (writing productions) becomes the current input. This concept of a constantly changing, interacting, and self-organising complex system is an essential characteristic of Dynamic System Theory.

Our study attempts to reveal the relative importance of the interconnected variables, input outside the classroom and vocabulary development. We will model the system in which the development is iterative, and the pattern is non-linear. In order to do so we first observed the developmental patterns in the data and decided that the logistic model was suitable for our study. Further research into the theoretical background of the model demonstrated that it was compatible with the concepts on which our data is based: iterativity, non-linearity, and time-dependency. Applying the logistic model to interpreting the trajectory of the data enables us to obtain the values of the three main variables, which meaningfully quantify the vocabulary development.

4.3.8 The modelling procedure

The general impression of the first look at the trajectory is usually fluctuant as there is a lot of variability. It is natural to find variability in the trajectory as there are many factors influencing the performance of the learners. However, to obtain a more clear idea of what the general trajectory looks like, it is necessary to remove some of the variability from the data sets, whilst at the same time ensuring there is no loss of crucial information from the data. One technique to remove the variability is data smoothing. We opted for one commonly accepted smoothing method- smoothing spline (Green & Silverman, 1994), which locally fits a cubic smoothing spline to the data sets and is more dynamic than creating linear trend lines. This smoothing method was carried out in “R” (http://www.r-project.org).

After obtaining the smoothed data, we proceed with the model fitting: to fit the logistic model with the smoothed data in order to obtain the values of the three variables. To find the best fitting, the smallest difference between the logistic model and the smoothed data must be ascertained. This difference usually refers to the smallest sum of squares. The values of the three variables are entered into the equation of the logistic model to generate the model trajectory, and the
difference between the model trajectory and the smoothed trajectory is calculated in terms of the sum of squares. This process is repeated until the smallest sum of squares is found. It is not feasible to manually enter the values of three variables thousands of times to find the least sum square. A curve-fitting program, Amoeba (Press et al, 1986) should be applied to execute the fitting task to automatically find the smallest sum of squares and then generate the values of the three variables.

4.4 Results

We intended to investigate how the lexical difficulty (represented by the measure of $V_{size}$) of writings was affected by different degrees of input outside the classroom. We first observed the trajectory of the data and the smoothed data of the $V_{size}$, then obtained the values of the three variables (initial value, learning rate, and the carrying capacity) in order to determine the logistic model, and then compared the mean $V_{size}$ for each degree (low, intermediate, high) of English input.

4.4.1 Developmental trajectory

The development of raw data is depicted in Figure 4.2. Skid’s data has more variability than the other three participants. Tina’s, Grace’s, and Gloria’s data has fewer fluctuations, indicating their writing performance was more stable than Skid’s. However, the trajectories of the data did not give us a clear picture of the development of $V_{size}$ due to the high variability and the scale. We obtained the smoothed data, leaving out some variability, to gain a clearer idea of how to explore the trend of the development.

4.4.2 Developmental trajectory of smoothed data

The trajectory of smoothed data is depicted in Figure 4.3 in the form of solid lines. Skid’s smoothed data is still more variable than the other three participants: Tina’s data shows an “s” curve pattern in the logistic model; Gloria’s and Grace’s data is identical, showing as an increasing line. We could conclude that the four learners in our study, judging from their writings, have increased their $V_{size}$ values over the five-month period.

4.4.3 Parameter values of the model fitting

By fitting the smoothed data with the logistic model, we were able to obtain the values of the three variables determining the development of $V_{size}$. We first
Input and vocabulary development

Figure 4.2: Trajectory of the raw measurements of lexical difficulty (V_size) of the four learners

had to ascertain whether the fitting was good enough to interpret the data. The trajectory of the fitted data (dashed line) is depicted in Figure 4.3 together with the smoothed data (solid line). As can be seen, the logistic model did not capture the entirety of Skid’s developmental pattern: there is great disparity between the trajectory of the smoothed data and that of the fitted data (the model). However, the logistic model fits well with the smoothed data of the other three participants.

The values of the three variables (initial value, learning rate, and carrying capacity) generated from the model fitting are displayed in Table 4.1. Since the smoothed data from Tina, Grace, and Gloria demonstrated a good fit with the model, we were able to use the logistic model to describe their lexical difficulty in terms of V_size. Tina had a lower initial value of V_size; she had a value of 2732 words to produce. However, she had a higher learning rate (0.119), and was therefore able to produce almost the same number of words (3846) as Gloria and Grace (3973 and 3778). Tina’s V_size increased when receiving increasing amounts of input outside the classroom (data point 1 to 17) but started to stagnate in episode 38, which was about the moment that the input outside the classroom became much less than in previous periods. We could conclude that Tina’s V_size increased with the higher degree of English input, but it also decreased with the lower degree of English input and that therefore, Tina was sensitive to the volume of input she was exposed to.

Gloria had a V_size value of 3531 words at the beginning of the experiment. These values increased with a lower learning rate (0.03) but steadily reached 3973 after the increasing exposure of input outside the classroom. Grace had a V_size value of 3070 words at the beginning. Her V_size grew with the same learning rate as
Input and vocabulary development

Figure 4.3: Trajectory of smoothed and fitted data of V_size of the four learners.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Initial Value</th>
<th>Learning Rate</th>
<th>Carrying Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skid</td>
<td>4307</td>
<td>1</td>
<td>5094</td>
</tr>
<tr>
<td>Tina</td>
<td>2732</td>
<td>0.119</td>
<td>3846</td>
</tr>
<tr>
<td>Gloria</td>
<td>3531</td>
<td>0.03</td>
<td>3973</td>
</tr>
<tr>
<td>Grace</td>
<td>3070</td>
<td>0.03</td>
<td>3778</td>
</tr>
</tbody>
</table>

Table 4.1: Values of the three variables for the four learners.

Gloria’s learning rate (0.03) and ended up as 3778 with the increasing volume of English input. We can conclude that Gloria and Grace were less sensitive to the volume of input outside the classroom than Tina.

Skid’s model showed a relatively poor fitting, as it did not fit well with the smoothed data. As can be observed from the Figure 4.3, the logistic model was able to describe the initial value accurately but failed to capture the learning rate and the carrying capacity since the development pattern of the data in the middle and the final stage largely deviated from the expected pattern of the logistic model. A very high learning rate was generated from the model fitting and the carrying capacity reached its final value at a very early stage. In other words, the logistic model generally did not reflect the reality in Skid’s data set. Therefore, we can only state that Skid had about a level of 4307 V_size at the beginning of the experiment.
4.4.4 Mean of vocabulary size in each input stage

We show the mean of V\_size in each input stage along with the timeline in Table 4.2 and 4.3. Skid’s V\_size, even during the period of low input, still increased through reflection and practice of the words, which he originally knew receptively but started to produce in his writing. Skid was highly motivated by the input from movies, where his V\_size grew about 1469 words in size. However, Skid was less motivated by the input from online articles and his V\_size decreased by 979 words.

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Number</td>
<td>1-19</td>
<td>20-38</td>
<td>39-56</td>
</tr>
<tr>
<td>Input Volume</td>
<td>Low</td>
<td>Intermediate Movie</td>
<td>Intermediate Reading</td>
</tr>
<tr>
<td>Mean of Vocabulary Size</td>
<td>4415</td>
<td>5884</td>
<td>4905</td>
</tr>
</tbody>
</table>

Table 4.2: Mean of Vocabulary Size in each input stage of Skid.

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Number</td>
<td>1-10</td>
<td>11-20</td>
<td>21-37</td>
<td>38-56</td>
</tr>
<tr>
<td>Input Volume</td>
<td>Low</td>
<td>Intermediate Reading</td>
<td>High</td>
<td>Intermediate Movie</td>
</tr>
<tr>
<td>Tina</td>
<td>3060</td>
<td>3270</td>
<td>4111</td>
<td>3652</td>
</tr>
<tr>
<td>Gloria</td>
<td>3620</td>
<td>3570</td>
<td>3958</td>
<td>3842</td>
</tr>
<tr>
<td>Grace</td>
<td>3030</td>
<td>3120</td>
<td>3600</td>
<td>3621</td>
</tr>
</tbody>
</table>

Table 4.3: Mean of Vocabulary Size in each input stage of Tina, Gloria, and Grace.

During the low input period, Tina’s V\_size did not grow, but it increased by 210 words when receiving intermediate input from reading online news and by 841 words when receiving high input. When receiving less input in the last period, her V\_size reduced by 459 words. Tina showed that she was sensitive to the volume of input: the more she was exposed to the English input, the more her V\_size increased. From her writings, we found that she liked to use new words gained from chats with friends in her writing, especially the more native-like chunks. Her writings reflected her preferred way of saying things while interacting with other people in English, because she liked to sound and speak like a native speaker.

During the low input and the intermediate input period, Gloria’s V\_size did not grow, but it increased by 388 words in the high input stage. Her V\_size reduced by 116 words in the last intermediate input stage when she received less input than the previous stage. Gloria only showed a slight growth in the high input period; she was less sensitive to the volume of input, showed a stable learning curve, and
carefully used new words gained from reading online news, especially the words of low frequency.

Grace’s V.size, even during the low input period, gradually increased, but her growth slowed down during the following intermediate input period. Her V.size increased by 480 when receiving high input, and she was able to retain her words in the following intermediate input period with a difference of only 21 words. On the other hand, the data suggested that Grace had a delayed effect from the exposure of input. Her V.size tended to grow, to stabilize, to grow again, and to stabilize again. From her writings, we found that she usually only used the words she already knew very well and only started to use new words when she was completely confident in how to use them correctly. This pattern may explain the delayed effect of the exposure of the input on V.size.

4.4.5 Summary of results

The first research question asked whether learners of English developed their writing in terms of lexical difficulty (represented by the measure of V.size) with more exposure to input outside of the classroom. The increase of lexical difficulty was determined by the difference between the mean of the lexical difficulty of low and intermediate input or between the mean of the lexical difficulty of low and high input. The four learners showed an increase in their lexical difficulty. When looking at the increase of the whole development, we found that the increase of lexical difficulty, determined by the difference between the carrying capacity (the final vocabulary size learners reached) and the initial value (the beginning vocabulary size learners held), increased by about five-hundred words in a period of five months with extra incidental English input, mostly from movies and reading. Since the three girls (Tina, Gloria, and Grace) attended merely traditional Taiwanese English class where hardly any English input existed and Skid did not have any English lessons, it is likely that this effect is caused by the extra enhanced input.

The second research question asked whether the learners of English decreased their lexical difficulty with less exposure to input outside of the classroom. The decrease of the lexical difficulty was determined by the difference between the mean of the lexical difficulty in the earlier higher input stage (high input stage) and the mean in the later lower input stage (low input stage). Tina, sensitive to the volume of input, demonstrated an immediate decrease in her lexical difficulty. Grace and Gloria’s lexical difficulty did not decrease immediately, as there may have been a delayed learning effect from the input on their writings.

The third research question asked whether the logistic model could meaningfully interpret the development of lexical difficulty for these four learners. We found that the model fitting was ideal in Tina’s, Gloria’s, and Grace’s data. The values of the three variables, initial value, learning rate, and carrying capacity, were sufficiently meaningful to describe the lexical development of these three learners. Skid’s data,
however, did not demonstrate a trajectory similar to the logistic model. We were only able to use the initial value to describe Skid’s beginning state of the lexical development.

### 4.5 Discussion

The objective of this study was to find out in what way input outside the classroom influences written production in the L2. The logistic model analysis showed that three of the participants in our study (Tina, Gloria, and Grace) developed their lexical difficulty with input outside the classroom. They did not have equivalent initial values of lexical difficulty but reached approximately the same degree of lexical difficulty by the end of the experiment due to different learning rates. We could argue that the differences in sensitivity to the volume of English input between the three different learners were determined by the learning rate: Tina’s vocabulary development was of a higher learning rate; she was more sensitive to input outside the classroom than Grace and Gloria, whose learning rates were lower. Skid’s lexical development did not develop in the same way as the other three participants. His vocabulary development showed considerable variability; in fact, his smoothed data indicated that there might be three overlapping logistic models. A single logistic model could not capture all of the characteristics of his lexical development.

The logistic model successfully quantified the development of lexical difficulty for Tina, Gloria, and Grace with the three variables: initial value, learning rate, and carrying capacity. The values of these three variables enabled us to make explicit observations of how input outside the classroom influences lexical development. This preliminary model of L2 lexical development was, however, not able to represent Skid’s data. We should consider at least two more variables in the equation of the logistic model: the quantity of input and the quality of input. When the quantity of the input is lowered to a certain level, the lexical development not only ceases to grow but also begins to decrease. Additionally, the quality of the input, which could potentially be determined by the absolute value of the difficulty of input and the level of learners’ vocabulary size, could influence the lexical development. These two variables have not yet been included in our current logistic model.

The variability of the data is large within each individual. Two factors may have led to the high variability. Firstly, we observe a development over time, where the increase of vocabulary is expected as there is input available. Variability usually takes place when there is a change in the language development (Spoelman & Verspoor, 2010). Secondly, the adult learner, Skid, showed very high variability within his data sets because he immediately used some low frequency words which he receptively acquired from input in his writing productions— even the writing topics were not relevant to the input learners were exposed to. As mentioned
previously, the value of $V_{\text{size}}$ is very sensitive to the words in the low frequency bands. There is therefore more variability to be found in Skid’s data.

The other factor leading to the high variability is the topic of the writing. Although all topics were general and did not require much specific vocabulary, the performance of the writings on different types of topics may depend on personal experience and preferences. This is an inherent challenge for all studies that use longitudinal approaches to investigating writings in a natural environment. Our study looked at the beginning learners of English, who have a limited number of lexical items, with only the focus on the topics related to their life experience. The effect of writing topic cannot be avoided, and we have tried to minimize it by giving rather simple types of writing topics.

The lexical difficulty increased by about five hundred in $V_{\text{size}}$ value through incidental English input over five months. The time window of the observation may not be long enough for learners to show a larger growth in lexical difficulty, especially when looking at the productive lexical difficulty. Learners may receptively know some words partially and then may have a more thorough understanding of the words through input or may activate the words through input. However, learners may also acquire some new words, yet may not be able to use them productively in a text. The time window of this study may not have been long enough to allow for incidental learning by simply watching movies and reading some online news. We will continue to collect the writings of these participants to extend the time scale of the observations.

4.6 Conclusion

This study has attempted to investigate how the development of L2 lexical difficulty develops with input outside the classroom. We have used the logistic model to account for the lexical development. This model quantifies the important components determining the lexical development, which are the initial value, the learning rate, and the carrying capacity. Though the preliminary logistic model was not able to describe the vocabulary development of one participant, the trajectory of this participant indicated two possible variables to add in the equation of the logistic model, the quantitative input, the total amount of input exposure, and the qualitative input, the degree to which learners absorb the vocabulary from the exposure.

Learners, with input outside the classroom, kept themselves in an English context and benefited from having more opportunities for using English. With pure incidental English input in the form of movies and online reading, the increase of lexical difficulty was approximately 500 words in five months. Although this longitudinal study has collected data on a relatively short-term time scale, it has explored the microscopic level of understanding the actual lexical development.
Our study attempts to complement the previous vocabulary learning studies by going beyond a static means of looking at the data to an understanding of the dynamics of vocabulary learning.