Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20). 2008. Volume 1.
Edited by Marjorie K.M. Chan and Hana Kang. Columbus, Ohio: The Ohio State University. Pages 309-326.

The Acquisition of Chinese Shape Classifiers By L2 Adult Learners

Neal Szu-Yen Liang
The University of Texas at Arlington

Historically, much attention has been given to the acquisition and the development of Chinese classifiers by L1 children. Little, if any, is known about how non-native adult speakers of Chinese acquire this linguistic feature. To that end, the current study aims to explore the acquisition of eight shape classifiers denoting one-, two- and three-dimensional objects by adult speakers of English and Korean with various Chinese proficiency levels. Their task was to match ten objects made of clay with one of eight phrases that best describes the denoted object. The findings show that 1) a positive relationship exists between subjects' Chinese levels and their performance in this task; 2) Korean subjects minimally outperformed their English counterparts ONLY at certain stages; 3) 2-dimensional classifiers are best learned followed by 1-dimensional and then 3-dimensional classifiers. Pedagogical implications based on these results are suggested.

1. Introduction

When learning a new language, linguistic features that are obligatory in the target language but not present in the source language tend to be a source of difficulties for learners' progress. For example, for Chinese learners of English, linguistic features such as tense conjugation, mass-count distinction, using definite articles are the root of errors even after knowing other grammatical rules perfectly for years. One such feature in Mandarin Chinese, the classifier (henceforward CL) system, falls into the category described above for speakers of Indo-European languages. Not only do L2 learners need to memorize the semantics of classifiers, they also need to understand the underlying principles of selecting correct classifiers that varies depending on the contexts. Such guiding principles are intuitive to native speakers but are not so explicable, nor so easily predictable, to L2 learners.

The following example illustrates such a Number + Classifier + Noun construction. It is obligatory that a classifier be inserted between a Number and Noun in modern Mandarin Chinese, e.g. 三隻雞 *sān zhī jī* 'three chickens':

三	隻	雞
sān	zhī	jī
Num	CL	N
Three	animal-like-thing	chicken
'three chickens'		

The classifier 隻 *zhī* here is used when denoting most animal-like things. Likewise, a classifier 張 *zhāng* is used when the referents are objects with a flat surface. Thus, 'three tables' is 三張桌子 *sān zhāng zhuōzi*, 'three paintings' is 三張畫 *sān zhāng huà*, 'three faces' is 三張臉 *sān zhāng liǎn* and so on. However, 張 *zhāng* also collocates with 'bow' 弓 *gōng*, 'mouth' 嘴 *zǔi*, 'Chinese zither' 琴 *qíng* and other items that do not have a characteristic of flatness inherently. While some of these collocations can be explained away by historical developments within classifier systems, others do not have a good explanation; therefore, speakers simply have to memorize such constructions by rote. Not all native speakers can give a clear explanation with much confidence, but they all have no problem using it. So, how do L2 adult learners of Chinese cope with such somewhat confusing and unpredictable phenomenon and to what extent do their learning patterns in this area differ from that of the development of L1 child learners?

To that end, the current study aims to describe, analyze and explain the acquisition of Chinese classifiers by L2 adult learners. Inspired by Hu (1993), Fang (1985) and others' studies, the present study employs a comprehension experiment to investigate the development and acquisition of Chinese classifiers denoting nouns of different shapes by fifty eight adult L2 learners of Mandarin Chinese. This study will explore 1) the relationship between subjects' performance on this task with their Chinese Proficiency Level; 2) the relative order of acquisition of different classifiers; 3) the extent to which native speakers of Korean and English differ in terms of acquiring Chinese classifiers; 4) whether subjects' progression patterns show similar trend across different types of shape classifiers or certain classifiers have unusual patter.

The motivation of the current study is to extend the scope of the current literature which emphasizes 1) the nature and construction of classifier systems across language families of the world and 2) the phenomena of classifier acquisition by L1 children learners. Particularly, the current literature provides ample data and findings about how L1 children acquire Chinese classifiers, but one can hardly find studies that deal with such phenomenon exhibited by L2 adult learners of Chinese. It is envisioned that the results of this study would thus have a pedagogical contribution to the field of second language acquisition, particularly to that of Chinese language acquisition.

The three major components in this article are: 1) a literature review that contains previous research relevant to the current study; 2) a second part familiarizing readers with the subjects, materials, procedures and coding/scoring involved in this experiment; 3) a third section providing the results and their implications.

2. Literature Review

Researchers believe that linguistic classifications 'may be reflexes of basic cognitive categories' (Adams and Conklin 1973: 1). Such an idea motivates the current study which explores the relationship between linguistic categorization and human cognitive categorization. In this section, studies about how both L1 and L2 learners acquire this linguistic feature will be discussed.

2.1. L1 Acquisition of Chinese Classifiers

Studies concerning L1 children's acquisition of Chinese classifiers examine children's understanding of different aspects of the classifiers and the developmental progress at various ages. These studies have yielded somewhat similar results in certain aspects while somewhat different findings in other aspects are reported as well. In terms of methodology, two types of studies are found: Erbaugh's (1982, 1984, 1986) longitudinal approach and the experimental approach by others (e.g. Fang 1985; Loke and Harrison 1986; Chang 1983; Carpenter 1991; Loke 1991; Hu 1993a, b). They usually concern issues such as a) the emergence order of classifiers denoting different semantic domains; b) the relation between L1 children's cognitive development and their classifier acquisition, and c) the process of acquisition. Some of the common findings include that a) L1 children have a solid knowledge of the basic syntactic structure of classifiers at a very early age; b) their acquisition of classifier vocabulary is very much delayed compared to noun acquisition; c) although the findings of the order of Chinese classifier acquisition are different, it is very common that children over-generalize the general classifier 個 *gė* as a 'syntactic place-holder' (cf. Fang 1985, Hu 1993a); and d) they are very conservative in using classifiers.

In analyzing how children acquire noun classifier systems, Erbaugh (1984) finds that shape plays a more influential and stable role in sorting than function. Later on, Erbaugh (1986) further specifies her findings of both developmental and historical trends regarding acquisition of Chinese classifiers by L1 children. They can be summarized as follows:

1. Valued items before common ones, both before conventionalized sets.
2. Discrete, countable, portable concrete objects before large immovable ones.
3. Measures before special nouns classifiers.
4. Unique reference before prototypical.
5. Abstraction by extension specially rigid, horizontal length, before on a plane. Small size more prominent and earlier than shape, though roundness and squareness become common. Large size unmarked.
6. Classifier with number before demonstrative, near before far. Both before the pro-form.

In terms of shape classifiers, data from Erbaugh (1986) and Fang (1985) show that children acquire classifiers denoting one dimension or length first, then two dimensions or flatness and three dimensions or roundness last. One of the critics of Erbaugh's work focuses on the inadequate number of subjects which thus makes the findings less representative. In contrast, others' findings show different results. For example, Hu (1993a) suggests that the order of shape classifiers acquisition is two dimensions first, one dimension second and finally three dimensions. Still another inconclusive result by Loke and Harrison (1986) shows that the order of shape classifiers acquisition is three dimensions first, then one dimension and lastly two dimensions. Their findings are also more in agreement with the perceptual saliency order of Clark (1977) and Andersen (1978), which states that three dimensions and smallness were most perceptually salient to children. In summary, no final conclusion has been drawn regarding the acquisition order of shape classifiers by L1 children.

The literature so far discussed focuses on L1 speakers' use of classifiers. We do not know whether, or the degree to which, the above findings can apply to L2 adult learners of Chinese classifiers acquisition. The following section briefly summarizes a study relevant to this area.

2.2. L2 Acquisition of Chinese Classifiers

Polio's (1994) work provided some insight into how L2 adult learners acquire Chinese classifiers. Her 21 English and 21 Japanese adult speakers learning Chinese in Taiwan first viewed a short film that contains narratives of invisible referents. They were then asked to tell the story in the film to a native speaker of Chinese. The tester then examined their use of classifiers and found that a) NNSs had no problem using a classifier in obligatory contexts; b) they often included too many classifiers which makes it ungrammatical; c) they did use special classifiers, but only occasionally; d) they were able to self-correct the mistakes; e) there were a few cases where NNSs used unacceptable special classifiers.

These findings answered some questions regarding how L2 learners of Chinese develop their understanding and using of Chinese classifier systems, however, much remains unexplored. As Polio herself points out "there is much potential for research in examining how second language learners classify referents in relation to how first language learners classify referents" (Polio 1994: 63). Based on the literature reviewed thus far, the current research is motivated in order to bring some insight into the understanding of how L2 learners acquire Chinese classifiers and potentially provide pedagogical benefits to both teachers and L2 adult learners of Chinese.

3. Methodology

In this section, the subjects involved in this study will first be introduced, followed by the materials and methods/procedures for eliciting data. Finally, I present a detailed explanation of data coding and scoring for analysis purposes.

3.1. Subjects

There we two major groups of subjects in this study: L2 adult learners of Chinese and adult native speakers of Chinese. The use of Chinese classifiers by the first group is the source of data analysis while that of second group's serves as a control for the conventional use of classifiers.

There were 58 L2 Adult Learners of Chinese, henceforward NNSs, participated in this experiment. They were divided into three groups according to their Chinese proficiency level, henceforth CPL, with the Advanced level in Group I, the Intermediate level in Group II and the Novice level in Group III. Factors determining the subjects' placement include highest academic level achieved in learning Chinese, length of time studying Chinese, NSs' assessment and others. Detailed explanation in this regard is given in Section 2.3.

In order to determine whether there is any difference in Chinese classifier acquisition between subjects whose native language is a classifier language and those whose native language is not, native speakers of Korean and English were recruited in this study. Since Korean is a classifier language and English is not, it is predicted that Korean subjects should perform better than their English counterparts when learning Chinese classifiers. The numbers of Korean and English speakers recruited are the same. The following table shows the numbers for the two groups with three Chinese proficiency levels.

	Novice	Intermediate	Advanced	Total
Korean	9	8	12	29
English	12	9	8	29
Total	21	17	20	58

Table 3.1: Distribution of NNS by 1st language and CPL

The criteria for selecting subjects of L2 learners were that the subject must:

1. be 18 or above;
2. be a speaker whose first language is not Chinese;
3. be a speaker whose first language is either Korean or English;
4. have studied Mandarin Chinese for at least one semester;
5. be able to count numbers at least from zero to ten in Mandarin Chinese;
6. have no known visual impairment;

If any one criterion described above was not met, he or she was excluded from further participation of the experiment. Also, a questionnaire was designed to gather subjects' information to determine their qualification for the experiment. There are other questions in the questionnaire that ask for subjects' personal, academic and social life, such as highest education, numbers of Chinese speaking friends, length of time living in a Chinese-speaking community and others that might be factors of determining their CPL. The recruitment of the subjects for this study was done mostly in Taiwan with only 4 exceptions which were done in the US. Those tested in Taiwan were students of Chinese summer programs offered by either Taiwan Normal University or Furen University. They were approached randomly and agreed to take part in the test voluntarily. The test lasted between 30 to 45 minutes and was conducted individually whenever the time allowed. Occasionally, there were groups of 2 or 3 subjects taking the test at the same time due to the time constraints. In such cases, subjects were instructed not to share their responses or be influenced by others' presence.

For the second group, ten adult native-speakers of Chinese, henceforward NSs, were recruited locally in the US. Their home language had to be Mandarin Chinese for them to be qualified for this experiment. Those speak other dialects, such as Cantonese or Hakka, at home will not be considered. All of them have received at least college education in Mandarin Chinese in their homeland. The use of Chinese classifiers by these adult NSs serves as the control data for this study. The test was held at the NSs' home and the instruction was given in Chinese.

Due to the differences in language use between speakers from mainland China and Taiwan, the selection criteria here excludes those speakers from the former area. There are at least two reasons for such exclusion: 1) with fewer variables to consider, the data would be more controlled and uniformed if they were drawn from one group of subjects with similar background; 2) since most of the NNSs were recruited in Taiwan, it justifies the selection of only NSs from Taiwan. This makes the comparison between the NNS and NS groups more compatible and meaningful.

3.2. Materials and Procedures

The study employs a comprehension experiment designed to elicit subjects' understanding of eight shape classifiers. This section describes the selection of these classifiers, the objects created for and the procedures of the experiment.

Inspired by Fang (1985), Hu (1993) and others, the goal of this experiment is to investigate NNSs' ability in understanding classifiers denoting various shapes, sizes, and textures, with more emphasis on shapes. The types of shapes can be divided into one dimension (條 *tiáo* and 根 *gēn*, for rigidness), two dimensions (片 *piàn* and 張 *zhāng* for flatness), and three dimensions (顆 *kē,* 糰 *tuán* and 粒 *lì* for roundness, and 塊 *kuài* for cube). There are subtle differences between classifiers within each dimension group which will be explained next.

In the one-dimensional group, 條 *tíao* and 根 *gēn* differ in rigidness. For example, 條 *tiáo* co-occurs with flexible objects such as 繩子 *shéngzi* 'rope' while 根 *gēn* collocates with objects that are more rigid such as 香蕉 *xiāngjiāo* 'banana'. In the two-dimensional group, 片 *piàn* denotes objects of irregular shape while 張 *zhāng* denotes objects with square or rectangular shapes. For instance, 一片樹葉 *yípiànshùyè* 'a leaf' vs. 一張紙 *yìzhāngzhǐ* 'a piece of paper' in which a leaf is irregular in shape while a piece of paper usually has a square or rectangular shape. For three dimensions, 顆 *kē* and 粒 *lì* are usually used with more solid and round objects whereas 糰 *tuán* co-occurs with objects that are mushy or squishy. Furthermore, 顆 *kē* denotes things that are bigger than those denoted by 粒 *lì*. Lastly, 塊 *kuài* denotes objects with a cubic shape. Thus, typical objects denoted by 顆 *kē,* 粒 *lì*, 糰 *tuán*, and 塊 *kuài* are 西瓜 *xīguā* 'watermelon' (big and round), 花生米 *huāshēngmǐ* 'peanut' (small and round), 麵糰 *miàntuán* 'dough' (round and mushy) and 蛋糕 *dàngāo* 'cake' (cubic). The goal of selecting these objects is to reveal whether the subjects understand the subtle differences among these classifiers in denoting objects with different shapes, sizes and textures. The differences between these classifiers and typical objects belonging to that group are listed in the following table:

Shape	**Classifier**	**Salient Feature**	**Typical Objects**
One-Dimension	條 *tiáo*	slender, flexible, bendable	snake, worm, rope, river, road, …
	根 *gēn*	rigid	banana, cigarette, stick, match…
Two-Dimensional	片 *piàn*	thin with irregular edges	leaf, land…
	張 *zhāng*	thin with regular edges	paper, ticket, picture, face, table....
Three-Dimensional	糰 *tuán*	round & irregular shapes; squishy	dough, cloud, animal dropping,…
	顆 *kē*	round & bigger	basketball, watermelon, rock, planet,…
	粒 *lì*	round & smaller	rice, sesame, sand, grain, marble,…
	塊 *kuài*	cubical	Ice cube, brick, cake, nugget,...

Table 3.2: Classifiers by dimensions with salient feature and typical objects they denote

With these eight classifiers selected, ten corresponding objects made of children's modeling clay were created. Eight of these ten objects were each made to fit one and only one characteristic denoted by the eight classifiers discussed above. For example, to represent a typical object denoted by the classifier 塊 *kuài* , a cubical object similar to a sugar cube was made. Of these ten objects, two were made as foils with random irregular shapes and dimensions which prevented subjects from guessing the answers by eliminating the objects already selected.

Other characteristics of these objects such as color, weight, and texture were kept minimally different. Subjects were told that these objects all have a same name: 黏土 *niántǔ* 'clay', despite having different shapes and sizes. Each object was assigned a number written on a small card which was placed directly above the assigned object.

Next, eight phrases were presented to the subjects and they were asked to match each phrase with one and only one object based on the classifier presented in each phrase. Following are two example phrases:

一 條 黏土 ……………………………………..( )
yí tíao niántǔ
one CL-long-slender clay

一 張 黏土 …………………………………………………( )
Yí zhāng niántǔ
one CL-flat clay

Only the Chinese phrases and the pronunciation of each character, if unknown to the subjects, were provided in the actual test. The semantic gloss of each character is presented here only for explanation purposes. The only variable for all these eight phrases is the use of different classifiers. The correct selection depends on the subjects' understanding of the classifiers provided and the ability to link such understanding to the salient physical features each object exhibits. After the selection was made by writing down the corresponding number in the parenthesis, the subjects were asked to rate their level of confidence about their selections. The scale ranges from 1 to 5 with 1 being least confident and 5 being most positive about the choice.

Subjects were allowed to answer the questions in a random order and there was no time limit for the entire test.

3.3. Coding and Scoring

The raw data were coded before data analysis. This section explains how both nominal and ordinal data were coded and the criteria and schemes involved. I will first describe the coding process for data extracted from the Background Survey. Next, I present the data coding and scoring procedures for the actual experiment.

Coding for subject background data was done by creating a spreadsheet file. It includes the following characteristics of each subject:

(1) Gender
(2) Age
(3) Birthplace
(4) First Language
(5) Other Language(s) Learned

(6) Length of Time Studying Chinese
(7) Time First Studied Chinese
(8) Hours of Studying Chinese per Week
(9) Parents' First Language
(10) Length of Time Living in a Chinese-speaking Community
(11) Self-appraisal of Current Chinese Level
(12) Frequency of Using Chinese Daily.

Each subject was given an ID with a combination of 3 alphabet letters and 3 digits. For example, in an ID of L2EF01, L2 stands for the subject being a L2 learner of Chinese; E stands for the subject's first language as 'English'; F represents subject's gender as 'Female'; 01 represents the sequential number randomly assigned to each subject within their own group.

The NNS group was further divided into 3 subgroups: Novice, Intermediate and Advanced learners of Chinese. Criteria involved in determining the subjects' CPL include the above listed criteria from (6) to (12). Under each criterion, each subject was assigned a number of either 1, 2, or 3 based on the following schemes.

For criterion (6) Length of Time Studying Chinese, the dividing points, based on natural gaps, are 12 months and 24 months. Therefore, those subjects with less than 12 months' length of time studying Chinese received 1 point. Those between 12 and 24 months received 2 points and those with 24 months or more received 3 points. With such dividing points, there are 26 subjects received 1 point, 13 of them received 2 points and the rest of 19 subjects received 3 points. The point each subject received under this criterion was classified as INDEX 1 as seen in Table 3.3. This and 6 other INDEX points will be totaled as the final points to determine the subjects' CPL.

Code	**LnthTmStdChns(Mth)**	**Index 1**	**TmFrstStdyChns**	**Index 2**
L2EF01	10	1	12.63	2
L2EF02	36	6	34.90	3

Table 3.3: Sample Coding for NNS's Length of Time Studying Chinese & Time First Studied Chinese

The INDEXT 2 resulted from the criterion (7) Time First Studied Chinese. The longer ago a subject began studying Chinese, the greater point of 1, 2, or 3 will be assigned to that subject. The dividing points here are 12 months and 36 months. As such, those began studying Chinese less than 12 months ago received 1 point, those between 12 to 36 months received 2 points and those started 36 months or earlier received 3 points. Of the 58 subjects, 18 of them received 1 point, 23 subjects received 2 points and 17 subjects received 3 points.

For the INDEX 3, the criterion (8) Hours Studied per Week was calculated. Those who studied less than 10 hours per week received 1 point, those between 10 and 20 hours received 2 points, those have more than 20 hours received 3 points. As a result, 25 subjects received 1 point, 17 subjects received 2 points and 16 subjects got 3 points.

For the INDEX 4, more weight was given to the criterion (9) Parent's First Language. Those with a parent whose first language is Chinese received 3 points and others received 0 points. These extra points were added because I felt, upon speaking with them, could speak Chinese very well, or at least could understand my Chinese perfectly. Although they might not have taken many Chinese courses, it is apparent that their parents have a great impact on their CPL. With such assumption, 6 subjects received 3 points.

Code	**HrsStdied/wk**	**Index 3**	**PrntChns?**	**Index 4**	**LngthLvngC-C (Mth)**	**Index 5**
L2EF01	10	1	N	0	12	3
L2EM01	28	3	Y	3	10	2

Table 3.4: Sample Coding for NNS's Hours Studied/Week, Parental Lang & Length of Time Living in Chinese Community

The criterion (10) Length of Time Living in a Chinese-speaking Community provides the baseline for the INDEX 5. Those who have or had lived in Chinese-speaking community for less than 5 months received 1 point, those between 5 to 11 months received 2 points, and those for 12 or more months received 3 points. Following this standard, there are 22 subjects received 1 point, 21 received 2 points and 15 received 3 points.

For the INDEXT 6, I examined criterion (11) Self-appraisal of Current Chinese Level in which three levels were reported: Novice, Intermediate, and Advanced. Assuming the assessment of their own CPL was fair and accurate, those assessed as Novice received 2 points, those as Intermediate received 4 points and 6 points for those reported as Advanced. Again, the decision to give more weight to this criterion is solely subjective based on personal observation: the subjects tend to underestimate their Chinese levels so that they can avoid potential embarrassment if their testing scores did not come out as desired. For those who reported as Intermediate or better, they tend to have a certain degree of confidence in their Chinese levels and thus should deserve more points. At the end, there were 23 Novice, 24 Intermediate and 11 Advanced subjects.

For the INDEX 7, I divided the group by examining criterion (12) Frequency of Using Chinese Daily which is also a self-report of the subjects' judgment on their daily Chinese usage frequency. There are four levels: those who reported as 'Occasionally' received 1 point, those as 'Sometimes' received 2 points, those as 'Very Often' received 3 points and those as 'All the Time' received 4 points. The distribution of such allocation is as follows: 15 subjects reported as 'Occasionally', 20 as 'Sometimes', 17 as 'Very Often', and 6 as 'All the Time'.

Code	Chinese Level	Index 6	HwOftnSpkChns	Index 7	Final Point	CPL
L2EF01	Intermediate	4	Very Often	3	15	2
L2EF02	Advanced	6	Very Often	3	22	3

Table 3.5: Sample Coding for NNS's Chinese Level, Frequency Using Chinese & CPL

Finally, with the INDEX 1 to 7 tabulated and accounted for, the scores from each INDEX were tallied as the Final Point for each subject. The final points range from 8 to 25. They were further roughly and evenly divided into three groups based on the following dividing points: those scored 13 or less points received 1 point as true NOVICE, those scored between 13 and 18 received 2 points as true INTERMEDIATE, and those scored 18 or more points received 3 points as true ADVANCED learners. With such dividing points, there are 20 novice, 18 intermediate and 20 advanced learners.

The coding, scoring and organization of the data are explained in this section. The data are stored in a spreadsheet file with several columns. The first column is the subjects' ID based on the unique information for each subject explained earlier.

ID	1-1D1	1-C	1-S	2-2D1	2-C	2-S
L2EF01	8	5	0	10	4	8
L2EF02	1	4	10	10	4	8

Table 3.6: Sample Coding of NNS Responses on the First Two Questions

There are eight questions in the experiment and each question has three columns created to store three types of information. The first column is named '1-1D1' which contains the original answers given by the subjects. The first '1' stands for 'Question #1'; '1D1' stands for '1-Dimensional Classifier Type #1'. Therefore, with such a naming convention, one knows that the column '6-2D2' contains subjects' selection that reflects their understanding/knowledge of a classifier appearing in Question #6 which is used to denote 2-dimensional objects Type #2. The following Table 3.7 shows the eight classifiers with their corresponding Chinese character/Pinyin, dimension-type category and the ordinal numbers assigned to them in the actual experiment.

Question No.	Character/Pinyin	Dimension-Type	Assigned No.
1	條/*tiáo*	1-1	1
2	張/*zhāng*	2-1	10
3	糰/*tuán*	3-1	2
4	根/*gēn*	1-2	8
5	塊/*kuài*	3-2	9
6	片/*piàn*	2-2	7
7	顆/*kē*	3-3	6
8	粒/*lì*	3-4	3
Ø	Ø	3-5	4
Ø	Ø	3-6	5

Table 3.7: CL Categorization and Assigned Number

The second column created for each question has to do with subjects' confidence levels regarding their selections. This confidence levels range from '1' being 'Absolutely Do Not Know' to '5' being 'Absolutely Sure'. The names for these columns are given using the following method. In column '2-C', '2' stands for the number 2 question; 'C' stands for their 'Confidence Level', as seen in Table 3.8.

ID	1-1D1	1-C	1-S	2-2D1	2-C	2-S
L2EF01	8	5	0	7	4	2
L2EF02	1	4	10	10	4	8

Table 3.8: Sample Coding of NNS Confidence Level and Sample Scoring

The third column contains subjects' scores for a particular question asked in the experiment, as seen in Table 3.8, the subject L2EF01 gets a score of 10 for the question No. 2. The derivation of the scores for all of the subjects' eight questions is somewhat complicated and will be discussed in more details now.

To determine this score, it was necessary to reference L1 subjects' selections for each question. Ten L1 subjects participated in this research experiments and their answers were transferred and coded. Based on these L1 subjects' selections, a decision was made as to how many points any given selection is worth. For example, for 1-1D1, all of the L1 subjects' selection is '1'. That gives this selection '1' a 10-point score and all other nine possible selections will be worth 0 points. Therefore, as seen in Table 3.8, L2EF02's selection was 1 and 10 points were given. For subject L2EF01, her selection was 8 and she therefore received 0 points.

For another example with column 2-2D1 in Table 3.8 there are 8 out of 10 L1 subjects who selected No. 10 while the other two subjects chose No. 7 as their correct answer. No other selection was made for this question by the L1 subjects. Therefore, for L2 subjects' selections of 2-2D1, if their selection is No. 10, they will receive a score of 8 points, e.g. L2EF02's score; they receive 2 points if their selection is No. 7, e.g. L2EF01's score; and they receive 0 point if their selections are any of the other eight possible selections.

Finally, as seen in Table 3.9, two columns were created to sum up the subjects' total points and their score percentage of total possible points. The names for these two columns are 'Final Point' and 'Percentage' respectively.

With the subjects' background and test results recorded, coded and scored, further analysis of their performance is presented and explained in the next section.

ID	**Final Point**	**Percentage**
L2EM02	23	31
L2EM03	18	24

Table 3.9: Sample Scoring of NNS Final Point and Percentage of Possible Score

4. Result and Analysis

This section presents the result and a series of statistical analyses of the coded data drawn from the experiment by the L2 subjects. The aims are to organize and describe the data so they become informative and eventually provide implications as to how a Chinese teacher should approach students with different backgrounds when teaching different types of Chinese shape classifiers.

4.1. Correlation between CPL and Performance

In Figure 4.1, the scattered plot shows the relationship between the subject's CPL and their experiment scores. Clearly, there is a positive relationship observed between these two variables. The higher Chinese proficiency levels they have, the better they perform. This trend is similar to the trend found in the literature that L1 children's performance correlates positively with their ages and cognitive development. The pedagogical implication we can draw from this result is that teachers should teach classifiers to beginners with greater efforts than to advanced learners.

Next, when the subject group is divided into two subgroups, English and Korean, the previously observed trend remains with some variation. As seen in Figure 4.2, it shows that 1) Korean Novice and Advanced subjects do better than their English counterparts, but the difference is quite minimal; 2) English Intermediate subjects unexpectedly outperformed their Korean counterparts. Since Korean is a classifier language and English is not, one would logically assume that Korean subjects should perform much better throughout the 3 levels. However, the data show otherwise. Not only did Korean intermediate subjects perform worse than their English counterparts, the other two groups outperformed their English counterparts only minimally. The differences were not significant: between novice Korean and English groups, the p-value was 0.4615; between advanced Korean and English groups, the p-value was 0.6452. For the intermediate Korean and English groups, the difference was quite significant with a p-value of 0.0027.

The cause(s) of this rather counterintuitive result remain to be discovered. One can assume that the fact that Korean subjects did not progress as much during intermediate stage was caused by the overconfidence of their ability in learning this linguistic feature. Since Korean is also a classifier language, perhaps they did not put in as much effort as their English counterparts did. On the other hand, the English subjects viewed it as a challenging task since only measure words, but not classifiers are found in English. Hence, it required extra time and effort in this area for them and consequently they not only caught up with their Korean counterparts but also outperformed them at the intermediate stage.

The implication one can draw from this result is that teachers should pay different amount of attention to students at various stages with different backgrounds. At novice and advanced stages, English students would need more attention. At the intermediate stage, it is the Korean students who need an extra push when learning Chinese classifiers.

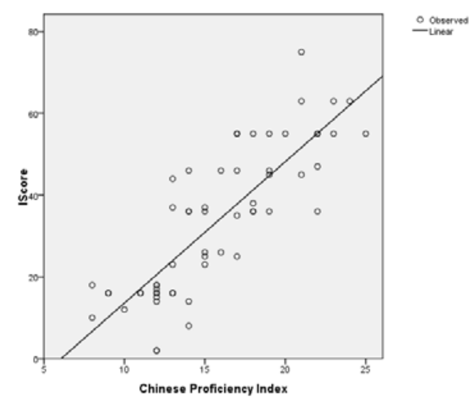

Figure 4.1: Correlation between CPL and Test Performance

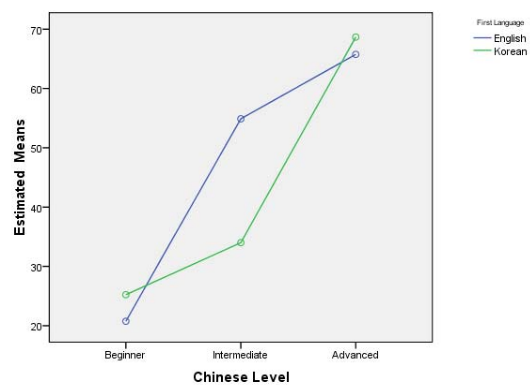

Figure 4.2: Estimated Means of NNS Performance Grouped by CPL

4.2. The Emergence Order

The emergence order of these three types of classifiers is the next issue to be explored. As seen in the Figure 3.3, 2-dimensional CLs are best learned by L2 learners followed by 1-dimensional and then 3-dimensional CLs. Notice the difference between 2-dimensional and 1-dimensional CLs is smaller than the difference between 1-dimensional and 3-dimensional CLs. This phenomenon is even clearer when only Korean subjects were examined. In Figure 3.4, Korean subjects’ performances on 2-dimensional and 1-dimensional CLs were relatively similar to their performance on 3-dimensional CLs.

When comparing the emergence order found in this study with those reported in L1 literature, one study has a similar result to the current study. Other studies on the emergence order show inconclusive results. For example, Erbaugh (1984) reported that emergence order for L1 children is 1-dimensional first followed by 2-dimensional and then finally 3-dimensional classifiers. On the other hand, Loke & Harrison's (1986) study claimed that 3-dimensional CLs were first learned followed by 1-dimensional and then 2-dimensional CLs. In yet another study, Hu (1993) showed that the emergence order is 2-dimensional first followed by 1-dimensional and then 3-dimensional CLs, which is similar to the current study's result.

The discrepancy might be due to the fact that the methodology designed and employed and the selection of subjects were quite different in these studies. For example, in Erbaugh's study, her sample size is rather small, only 4 children were included and their ages were relatively young as well (1;2 to 3;1). Based on findings by other studies, even by age three, children use very few special classifiers. With this limited subject pool, comparing results from this study to other studies is then not so compatible. Also, her methodology is the only one that is longitudinal. This factor is very likely the source that contributed the discrepancies.

In Loke & Harrison's study, on the other hand, their subjects' ages were older than the previous study and the sample size was greater as well. However, in their analysis, they considered the classifier 個 *ge* not just a general CL but also a 3-dimensional CL. They argued that in addition to being the classifier for humans and abstract entities, 個 *ge* 'can classify or reclassify *only* 3-dimensional objects since its classification of 1-dimensional and 2-dimensional objects would normally considered unacceptable or inappropriate by Mandarin speakers' (Loke & Harrison, 1986: 126). Consequently, it's not surprising to see their results showed that 3-dimensional CLs were learned first followed by 1-dimensional and then by 2-dimensional CLs.

In yet another format, Hu's sample size and method were more comparable to the current study. There were 24 subject L1 children in her experiment in which subjects' understandings of three types of dimension shape CLs was examined. With a similar methodology and sample size employed to the current study, her findings, not surprisingly, were more similar to mine than the others.

If the mean scores were grouped by NNSs' native languages, as seen in Figure 4.4, the above discussed emergence order is even clearer with English group than with Korean group. The chart can be summarized as 1). English L2 subjects performed better on 2D CLs but slightly poorer on 1D & 3D CLs than Korean subjects did; and 2). For both groups, 2-D CLs were best learned then 1-D and then 3-D CLs. The implications are that 1). teachers should assist English students more when learning 1-D & 3-D CLs and more with Korean students when teaching 2-D CLs; and 2) teachers should put more efforts on teaching 3-D CLs, followed by 1-D and then 2-D CLs.

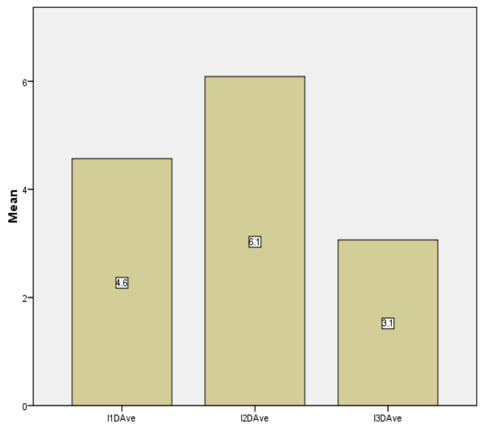

Figure 4.3: NNS Mean Scores Grouped by CL Dimension

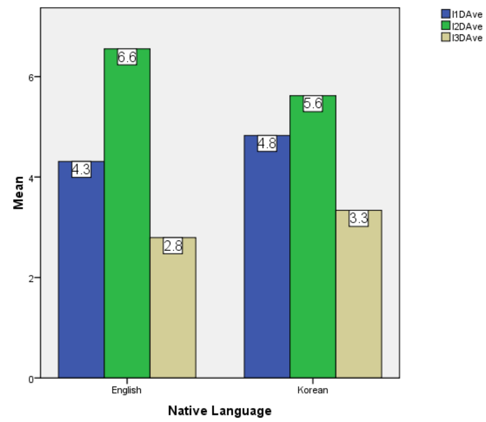

Figure 4.4: NNS Mean Scores Grouped by CL Dimension and 1st Language

4.3. Progression Patterns of Different Shape Classifiers

Finally, if we look at how subjects with different Chinese levels perform on different shapes of classifiers, something unusual appears. As can be seen in Figure 3.5, the progression of learning 1-D and 3-D CLs is pretty straightforward: the performance increases as their CPL advances. However, the progression of learning 2-D CLs is not so logical. The performance progressed initially as they moved from novice to intermediate stages. However, such progress turned downward as their CPL advanced from intermediate to advanced level. The cause(s) of such rather illogical phenomena and the conditions under which this downward U-shaped curve might occur remains to be investigated.

There are many studies that showed learners exhibit this kind of U-shaped learning pattern when learning other types of linguistic units. For instance, Abrahamsson (2003) investigated the relation between consonant deletion and vowel epenthesis in the development of word-final codas in Chinese-Swedish interlanguage. He found that his subjects' acquisition of Swedish codas 'exhibited relatively high accuracy rates at early stages, lower accuracy rates at later stages, and again high accuracy rates at more advanced stage'. Unfortunately, possible causes of this phenomenon were not provided. One of my ongoing research goals is to examine more closely these CLs and test them against other extralinguistic factors that might contribute to causes of these phenomena. Hopefully, this effect can provide suggestion as to what can be done by teachers to assist students to better learn the 2-D CLs.

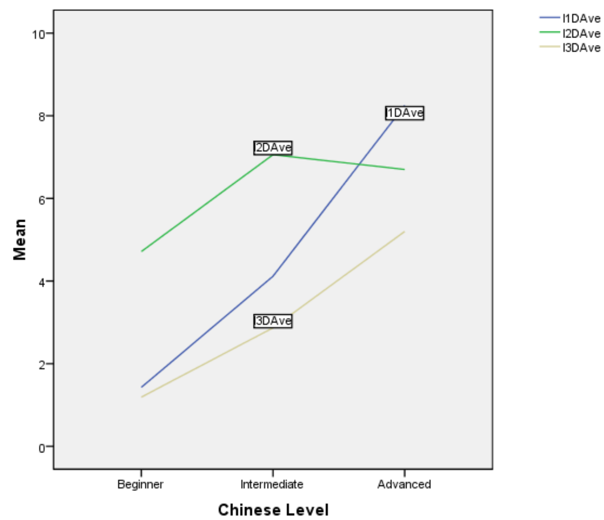

Figure 4.5: Correlation b/t NNS Performance and CPL Grouped by CL Dimension

5. Summary and Conclusion

Classifiers are said to have a close correlation with the ecological-Cultural environment of the speech community (Denny, 1976). This correlation is made clear by looking at the relationship between nouns and classifiers: ‘nouns have more to do with what is out there in the world, and classifiers more to do with how humans interact with the world (Denny 1976)’. The current research aims to provide some insight into how L2 adult learners of Chinese acquire Chinese shape classifiers and suggests approaches that language teachers should be aware of when teaching this linguistic feature to students with different L1 background. Based on research methodology used in L1 acquisition of Chinese classifiers, the results of this study show that 1) there is a strong positive correlation between subjects’ performance on Chinese shape classifiers comprehension test and their Chinese proficiency levels; 2) contrary to the prediction, Korean subjects outperformed their English counterparts only minimally at novice and advanced levels and were outperformed by their English counterparts at intermediate stage; 3) 2-dimensional classifiers were best learned by both subject groups followed by 1-dimensional and then 3-dimensional classifiers; 4) subjects showed an increasing progression as they advance from novice to advanced level when learning 1- and 3-dimensional classifiers. However, when learning 2-dimensional classifiers, their progress showed a reversed U-shaped progression pattern. The results from this study shed some light on how L2 adult learners of Chinese acquire Chinese shape classifiers which is relatively scantly discussed in the literature. Other types of experiments, such as classifier production test, and considering extralinguistic factors in identifying learners’ strength and weakness in learning this linguistic feature will be reported in my ongoing research projects.

REFERENCES

Abrahamsson, Niclas. 2003. Development and Recoverability of L2 Codas: A Longitudinal Study of Chinese-Swedish Interphonology. Studies in Second Language Acquisition 25.313-49.

Adams, Karen L. & Conklin, Nancy F. 1973. Toward a theory of natural classification.Papers from the ninth regional meeting of the Chicago Linguistic Society 1-10.

Andersen, Elaine S. 1978. Lexical universals of body-part terminology. In J. H. Greenberg (Ed.), Universals of human languages 3.335-68. California: Stanford University Press.

Carpenter, K. 1991. Later rather than sooner: extralinguistic categories in the acquisition of Thai classifiers. Journal of Child Language 18.93-113.

Chang, H.W. 1983. Pre-schooler's use of classifiers in Mandarin Chinese. National Taiwan University Press.

Clark, Eve V. 1977. Universal categories: On the semantics of classifiers and children's early word meaning. In Linguistic Studies offered to Joseph Greenberg on the occasion of his sixtieth birthday, ed. by A. Juilland, 449-62. Saratoga, CA: Anma Libri.

Denny, J. Peter. 1976. What are noun classifiers good for? Papers from the 12th Regional Meeting of Chicago Linguistic Society. Chicago Linguistic Society 122-32.

Erbaugh, Mary S. 1984. Scissor, Paper, Stone: Perceptual Foundations of Noun Classifier Systems. Papers and reports on child language development 23.41-49.

Erbaugh, Mary S. 1986. Taking Stock: the Development of Chinese Noun Classifiers Historically and in Young Children. In Craig (ed.) Noun Classifiers and Categorization 7.399-436.

Fang, Fuxi. 1985. An experiment on the use of classifiers by 4- to 6-year olds. Acta Psychologica-Sinica 17.4.

Hu, Qian. 1993a. The acquisition of Chinese classifiers by young Mandarin-speaking children. Boston, MA: Boston University dissertation.

Hu, Qian. 1993b. Overextension of animacy in Chinese classifier acquisition. In E. Clark (ed.). The Proceedings of the Twenty-fifth Annual Child Language Research Forum, 127-136. Stanford, CA: Center for the Study of Language and Information.

Loke, Kit Ken and Godfrey Harrison. 1986. Young children's use of Chinese (Cantonese and Mandarin) Sortal classifiers. In H. S. R. Kao and R. Hoosain (eds.), Linguistics, Psychology, and the Chinese Language, 125-146. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Loke, Kit Ken. 1991. A semantic analysis of young children's use of Mandarin shape classifiers. In A. Kwan-Terry (ed.), Child Language Development in Singapore and Malaysia, 98-116. Singapore: Singapore University Press.

Polio, Charlene. 1994. Non-native speakers' use of nominal classifiers in Mandarin Chinese. Journal of the Chinese Language Teachers Association 29.3, 51-66.