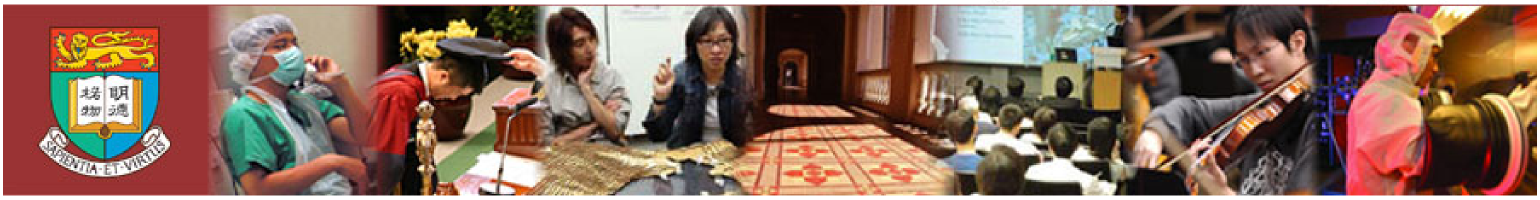

Title	General rule learning in school children
Author(s)	Lee, Oi-ting; 李靄庭
Citation	Lee, O. [李靄庭]. (2013). General rule learning in school children. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2013
URL	http://hdl.handle.net/10722/238518
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

General Rule Learning in School Children

Lee Oi Ting

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2013.

Abstract

The study investigated Chinese school age children's rule learning ability of non-adjacent dependency pattern with musical tones. Thirty-nine students with a mean age of 10 years and 5 months old were randomly assigned into 1 of 2 rule groups (*SD*= 10 months). They listened to musical tone streams generated according to non-adjacent dependency rule patterns. Afterwards, they judged if testing stimuli resembled the rules in training phase. Results showed that the participants were not able to extract non-adjacent rules from musical tones. Instead, the participants performed the task by learning adjacent dependency rule as the default. However, we cannot conclude there is no non-adjacent rule learning in Chinese population because the design of the study may be too difficult for school-aged children. For example, the variability of intervening elements was not sufficient and the differences on musical tones were very subtle for identification. Further research can consider the use of perceptual cues or melodies to investigate rule learning ability of non-adjacent dependency of nonspeech domain in Chinese population.

Over decades, the process of language acquisition is debatable among the developmental psychologists and linguists. Chomsky (1965) proposed the universal grammar. Infants are born with an ability to explore their native language with respect to the commonalty among all languages. In the late nineteen nineties, developmental researchers proposed new approaches towards language learning. One approach is to study the statistical learning ability (Aslin, Saffran & Newport, 1998; Saffran, Aslin, & Newport, 1996; Saffran, Newport & Aslin, 1996) and the other approach supports rule learning ability (Gómez & Gerken, 1999; Marcus, Vijayan, Bandi Rao & Vishton, 1999). These abilities are proposed to contribute to the syntax and vocabulary development (Marcus, 2000).

Statistical learning states that the calculation of co-occurrence of adjacent syllable pairs, which is reflected by transitional probability, facilitates speech segmentation in infants (Saffran et al., 1996). The speech sounds within a word will co-occur at a higher chance than that between words. As a result, low-probability adjacent pairs are likely to formulate word boundaries. Taking *happy#baby* as an example, the chance of co-occurrence of the syllable pair '*hap'* and '*py',* is greater than the pair '*py*' and '*ba*'. It is because there are a large variety of words possibly follow 'happy', such as happy boy, happy woman and happy dinner, etc. The calculation of the transitional probability of the adjacent syllable pairs allows the infants to segment '*happybaby'* into two words '*happy baby'*. Saffran and colleagues (1996) examined if eight-month-old infants were able to discriminate high probability syllable pairs against the low probability one in speech strings generated. The infants listened to nonstop streams of four trisyllabic artificial words repeatedly in the training phase. Then they listened to strings of familiar and unfamiliar sequence of those syllables. The infants showed more interests, determined by head turning preference (Kemler Nelson et al., 1995), when they listened to strings of syllable pairs with novel order. This suggested that the infants were able to fragment the nonstop auditory streams into meaningful units by statistical learning on consecutive syllables.

Another approach supports a higher order of language learning that requires rule abstraction and generalization of patterns into novel linguistic environment (for review, Gómez & Gerken, 2000). The study done by Marcus et al. (1999) examined if the infants were able to extract the adjacent dependency rule, which is defined as the rule pattern among consecutive syllables. Seven-month-old infants were exposed to syllable sequences either in the form of ABB, ABA or AAB, such as *ga ti ti, ga ti ga or ga ga ti*. The infants were first trained with one pattern (e.g. *ga ti ti*) and tested if they were able to identify this pattern in different syllables (e.g., *li na na*) against a novel pattern (e.g. *li na li*). The infants attended to the strings of unfamiliar pattern longer, which implied they were able to differentiate the trained pattern against untrained one. Since the testing speech strings were of novel syllabics following the same rule, the experiment demonstrated the acquisition of rule patterns from speech sequences. Other research carried out on older infants (Gómez & Gerken, 1999) also provided evidence on rule learning ability.

Marcus, Fernades and Johnson (2007) found that infants were not able to learn rules from musical tones, but the performance improved when there was prior exposure to linguistic stimuli. They proposed that rule learning in infants was specific to linguistic information. However, other researchers showed that infants were capable of extracting adjacent rule regularity among nonspeech stimuli, such as visual images (Johnson et al., 2009; Saffran, Pollak, Seibel, & Shkolnik, 2007). This raises the question whether rule learning ability is general across different domains, such as auditory, visual and linguistic materials, or specific to linguistic stimuli only. In our study, we will investigate the rule learning ability of nonlinguistic stimuli, musical tones, in school-aged students to revise the question if rule learning is domain specific.

Apart from adjacent dependency, the non-adjacent dependency is also important for syntax acquisition for some natural languages. Non-adjacent dependency may be more challenging as it requires the learning of nonadjacent items over an intervening elements in-between. In English, it is frequent in auxiliaries (e.g. is doing), inflectional morphemes (e.g. have talked), and number agreement (e.g. boy likes, the cars in the park are). In Chinese, about 40-50% dependencies are non-adjacent (Liu, Shao & Li, 2009). For example, the non-adjacent dependency can be found between temporal adverbs and the aspect markers (e.g. 曾經做過/cang4 ging1 zou6 gwo3/ 'has done'; 已經食了/ji2 ging1 sik6 liu5/ 'have already eaten') according to Ding (2008) and Xia (2011). Anaphoric reference demonstrates the non-adjacent dependency as well (e.g. 小明打緊波。佢好開心。/siu2 ming4 daa2 gan2 bo1. keoi5 hou2 hoi1 sum1/ 'SiuMing is playing a ball. He is happy.'). The rule learning ability of non-adjacent dependency was investigated in two studies (Gómez, 2002; Gómez & Maye, 2005). In experiments of Gómez (2002), forty-eight adults were exposed to one of two artificial languages (Language One: aXd, bXe, cXf; Language Two: aXe, bXf, cXd). A, b, c, d, e and f were fixed nonwords whereas X represented an artificial disyllabic word chosen from a set pool of nonwords. Pauses were inserted to signify word boundaries. The first two elements in both languages were the same and therefore, two languages differed in the last elements of strings (e.g. to differentiate aXd against aXe). The adults first listened to training stimuli from either one language for approximately 18 minutes. Afterwards they judged if testing stimuli followed the rule acquired in the training. Infants were tested with a less complex language using the same material (Language One: aXd, bXe; Language Two: aXe, bXd). Both infants and adults were able to abstract the non-adjacent dependency when the middle element X was highly variable. Gómez (2002) suggested the high variability of middle element X made the rule unpredictable according to consecutive syllables, i.e., the adjacent dependency pattern. Therefore, the participants switched their focus from adjacent rule to non-adjacent rule.

As both adjacent and non-adjacent dependencies are common in Chinese language, it is important to study the acquisition of these dependencies in Chinese population. In this way, the study evaluated if rule learning ability is universal among population speaking different languages. In addition, this research also investigated if rule learning ability is general across the nonspeech domain. The study adopted Gómez's (2002) design with similar language rules tested (i.e. Rule One: aXc, bXd and Rule Two: aXd, bXc). In our study, we substituted musical tones for the artificial nonwords and we replaced the middle disyllabic words X with two musical tones. Pauses were inserted between the two musical tones of middle elements X to simulate the syllabic nature of Chinese words, given the X was disyllabic. Therefore, each string was of four separate units in our study whereas each string was of three units in Gómez's study (2002). Instead of testing adults and infants, the study targeted children of ages between 9 and 13. We reasoned that children of this age range could perform the adult version of Gómez's (2002) rule learning tasks. Given that no research tested rule-learning ability of school-aged children, this study was the first to show whether school aged children were capable of abstracting non-adjacent dependency rules from nonlinguistic stimuli.

Creel, Newport and Aslin (2004) suggested learning of adjacent dependency and non-adjacent dependency occurred at the same time. Thus, we expected learning of both dependencies could take place. The two testing rules consisted of the same initial and middle element but only differed in the final element. If the participants can better differentiate trained rules against the untrained rules (e.g. aXc from aXd and bXd from bXc), they should be sensitive to the non-adjacent dependency. In order to show that the participants acquired the non-adjacent dependency instead of performing the task by tracing adjacent dependency or memorizing the whole string, we added a generalized condition. The generalized strings are composed of new middle element X, which the participants does not encounter in the training phase. If the participants can trace the non-adjacent dependency, they should be able to better differentiate the items of trained rules in generalized conditions against the untrained rules as well. The first element and the X elements in training co-occur at a higher chance than the first element and novel X elements. Therefore, if the participants traces the adjacent dependency pattern or memorizes the strings, the participants may better identify training items but not the generalized items.

Individuals with grammar impairments showed difficulties in extracting non-adjacent dependency rules (Hsu & Bishop, 2010). Grunow, Sapulding, Gómez and Plante (2006) showed that adults with language disabilities were not able to learn and generalize the non-adjacent dependencies even when there was high variability of the middle intervening element. In light of this, Hong Kong Cantonese Oral Language Scale (HKCOLAS) (T'sou, Lee, Tung, Man, Chan & To, 2006) was administered to reflect the grammar ability of participants.

Methods

Participants

Thirty-nine students (22 females and 17 males) ranging from 8 years 11 months to 12 years 11 months old (*M* = 10 years 5 months; *SD*= 10 months) were recruited in the study. They were randomly assigned into either one of two groups: Rule One group and Rule Two group, with 19 participants and 20 participants respectively. The data of three other participants tested were excluded because they scored -1.25SD below mean in HKCOLAS grammar subtests (T'sou et al., 2006). This avoided the impact of possible grammar impairments on rule learning ability (Hsu & Bishop, 2009). Faculty Research Ethics Committee at the University of Hong Kong approved the ethics.

Stimuli

There were two phases in the experiment, namely a training phase and a testing phase. During the training phase, each individual listened to auditory musical tones from either one of the two rules. There were two dependencies in each rule. Rule One (R1) was in the form of aXc, bXd and Rule Two (R2) took the form of aXd, bXc whereas a, b, c and d were definite tones. The X represented an element chosen from a pool of 24 tone sequences. We selected set pool size of 24 as high variability of intervening X element facilitated learning of non-adjacent dependency according to Gómez (2002). All piano notes were generated by

MidiPiano and the musical tone strings were adjusted and sequenced by Cool Edit Pro 2.0. Each musical note lasted for 500 ms and the tones were of the same loudness so that the duration or volume would not stand out any musical tones. There was 250 ms pause between tones and 750ms pause between each tone string, such that participants were able to differentiate musical tones within and between strings. Eprime version 1.2 randomized all stimuli and ensured smooth running of stimuli in both training phase and testing phase.

Training. All tones were selected from the same octave, which meant the musical notes lied between the middle C and middle B. Musical tones from other octaves were not selected as tones from different octaves gave perceptual cue to participants (Creel et al., 2004). A, b, c and d represented musical note C, D#, G# and B respectively. The 24 X tone sequences in training materials chosen were C#F#, C#A, C#A#, DE, DF, DG, ED, EF#, EC#, FC#, FD, FA#, F#A#, F#G, F#F, GE, GF, GA#, AF#, AF, AA#, A#E, A#F, A#C#.

Testing. There were four conditions (the grammaticality) in the testing phase, namely the trained, untrained, trained generalized and untrained generalized conditions with four strings in each condition. Eight strings were of the same rule as training phase (trained rule) and eight were from the other rule (untrained rule). For the trained and the untrained conditions, the middle elements were chosen from the presented 24 X tones sequences. For the trained generalized and the untrained generalized condition, the middle elements were two novel X tone sequences, which were FG and DA. Table 1 presented the testing strings in the experiment.

Table 1

Rule Dependency and Test Strings in the Experiment

	R1	R2
Dependency	a X c b X d	a X d b X c
Trained rule	C A# E G# D# A# E B C C# F# G# D# C# F# B	C A# E B D# A# E G# C C# F# B D# C# F# G#
Trained rule in generalized condition	C F G G# D# F G B C D A G# D# D A B	C F G B C D A B D# D A G# D# F G G#
Untrained rules	C A# E B D# A# E G# C C# F# B D# C# F# G#	C A# E G# D# A#E B C C# F# G# D# C# F# B
Untrained rules in generalized condition	C F G B D# F G G# C D A B D# D A G#	C F G G# D# F G B C D A G# D# D A B

Procedures

Hearing screening, grammar subtests of HKCOLAS and a musical tone judgment experiment were conducted for every participant. The order of three tasks was randomized among all participants. The hearing screening was administered to screen out hearing impaired participants as all the tasks relied on participant's hearing ability. The participants were screened at 1000 Hz, 2000Hz, 4000Hz and 500 Hz if the noise level was lower than 41.5 dB SPA (ANSI, 1991), measured by a sound level meter (model: TES 1350A).

We carried out a pilot to ensure smooth administration of tasks and to examine all the training and testing materials. Two participants were recruited and both failed to acquire the rules by exposing to 96 strings in the training phase (2 times X 2 dependencies X 24 tone sequences). Thus, in the actual study, the participants were exposed to a sum of 144 strings (3 times X 2 dependencies X 24 tone sequences) in training to familiarize them with the rule. Before training, they were informed they would be exposed to musical tones for about 10 minutes, which followed by a short test. After training, they learnt that musical tones they heard were generated according to a rule pattern, which was related to the musical tone order. The participants needed to judge whether the testing strings followed the same rule as those in training phase or not. If it followed the same rule, they were instructed to press 'yes' button. If the musical strings violated the rule in training phase, they needed to press 'no' button. Finally, they were asked to tell the rule explicitly after testing.

Results

All participants passed the hearing screening at 1000Hz, 2000Hz and 4000Hz at a noise level of about 50dB SPA in average. Table 2 presented the number of endorsement for grammaticality of two rules. Two way repeated measure analysis of variance (ANOVA) was carried out, with the rules (R1 and R2) as between subject variables and grammaticality (trained, untrained, trained-generalized and untrained-generalized condition) as the within subject factor. There was no significant main effects of grammaticality, $F(3,35) = 2.267$, $p = .098$, *ns* and rules, $F(1,37) = .213$, $p = .647$, *ns*. There was no significant interaction effect, $F(3,35) = .205$, $p = .892$, *ns*. The data for the two rules was then pooled because there was no significant difference between the two rules. Pairwise comparison showed there was significant difference between the trained and the trained generalized condition, $t(38) = 2.485$, $p = .017$, $< .05$, and also between the trained and the untrained generalized condition, $t(38) = 2.104$, $p = .042$, $< .05$. Comparison of other conditions did not show any significant difference.

The participants were asked to tell the rules explicitly after testing. Table 3 presented the summary of answers. Only a few were able to told one of the rules. Most participants judged either by guessing, by memory or by melody of musical strings.

Table 2

Number of Endorsements for Grammaticality

	Grammaticality							
	Trained		Untrained		Trained Generalized		Untrained Generalized	
	M	SD	M	SD	M	SD	M	SD
Rule 1	2.89	1.15	2.64	1.21	2.36	1.11	2.32	1.11
Rule 2	2.85	0.99	2.80	1.06	2.50	0.89	2.55	1.28

Table 3

Responses of Participants Concerning the Rules Identified

Responses	Number of participants	Percentage (%)
The participants were able to tell one of the rules.	3	7.7
The participants reported the untrained rules.	2	5.1
The participant was able to identify the difference in the last element.	1	2.6
The participants judged by melodies.	7	17.9
The participants answered by memory.	7	17.9
The participants thought the rule was 4 musical tones in a string.	5	12.8
The participants judged by guessing.	10	25.6
The participants thought they listened to all strings.	4	10.3

Discussion

The participants were not able to extract non-adjacent dependency rule from musical tones for both rules, given that the first element highly predicted the final element. Instead, they differentiated the trained tone strings against the novel tone strings, the generalized conditions.

Adjacent dependency versus non-adjacent dependency

The participants failed to differentiate the strings in the trained condition against the untrained condition. They were insensitive to the non-adjacent dependency embedded and possibly performed the task by tracking the adjacent dependency of the musical strings. Gómez (2002) suggested the participants tracked adjacent dependency as the default and they would only focus on other rule if they can better track any other regularity from the materials. Although we have 24 possible tone sequences for the middle element X, the participants might regard the middle element X as two separate elements since pauses were inserted in-between. If this were the case, we would have to view each X as two separate elements: X_1 and X_2. Since only eight different musical notes followed the initial element, the variability in X_1 was in fact eight only. The low variability of intervening elements may not be sufficient for the participants to shift their focus from adjacent dependency to non-adjacent dependency (Gómez, 2002; Gómez and Maye, 2005; Onnis, Christiansen, Chater & Gómez, 2003; Onnis, Monaghan, Christiansen & Chater, 2005). In addition, the use of nonlinguistic posed constraints on the acquisition of non-adjacent dependency (read Creel et al., 2004; Gebhart, Newport & Aslin, 2009; Marcus et al., 2007 for how difficult rule learning or statistical learning of nonspeech stimuli was). Without auditory perceptual cue, which group non-adjacent items with similar pitch or timbre, adjacent dependency took the priority over non-adjacent dependency (Creel et al., 2004).

From the result, the participants better discriminated the trained items against the generalized items, which reflected their focus was on the middle element X instead of the non-adjacent dependency. There are two possible explanations. First, the participants might reject the generalized X items by memory. However, this is of low possibility because the participants were exposed to each tone sequence with limited times (6 out of 144 times). Second, the calculation of transitional probability of consecutive tones can help differentiate tone sequences of trained order from the untrained one (Saffran, Johnson, Aslin & Newport, 1999). In our study, the participants might trace the adjacent dependency between the second and third element, i.e. between X_1 and X_2. The two X_2 in the generalized items were G and A which were infrequently used in the X_2 position of the training items (i.e., G: 2/24 times; A: 1/24 times). The participants in our study might reject the generalized items by calculating the chance of occurrence of the pair X_1 and X_2. Therefore, future studies should be more careful in selecting the middle element X as it seems to play an important role in participants' learning of non-adjacent dependency.

Learning musical tones was complicated

To perform the task, the participants were required to identify the absolute tone (i.e., the exact pitch) of the first and last element for abstraction of the non-adjacent dependency rule. To avoid giving auditory cues in the experiment, musical notes from the same octave were selected. As a result, musical tones differed by only three semitones were chose (i.e. C and D# for the first element; G# and B for the last element). However, the differences between C and D# or between G# and B may be perceptually quite subtle for identification, for example, they were required to distinguish C A# E G# from C A# E B. Benguerel and Westdal (1991) proved that individuals were not able to memorize every single musical note without any contexts. Only absolute pitch possessor was able to identify absolute musical note (Ward, 1999). It would definitely challenge the participants if they needed to memorize the final elements during the training phase and then, to distinguish them against the final elements of untrained rules in testing phase. In addition, individuals recognized musical tones according to intervals and contours instead of the exact pitch value (Schellenbery & Trehub, 2003). The recognition of random melodies was poor, which contributed to the fair ability to identify the dependences among tones (Saffran, 2003). The tone strings composed were of no melodic value as well, which made it more difficult to identify the absolute first and last musical tones of the two dependencies. These may explain the participants' inability to learn the non-adjacent dependency pattern in the experiment, for example, about two third participants (43.6%) claimed to perform the task by random guessing or rote memory. One tenth of participants (10.3%) even stated they have heard all musical streams in the test, which reflected the differences on final elements were too slight. A small number of participants (7.7%) were able to tell one of the two rules in the experiment. This may provide limited evidence on some learning of non-adjacent dependency pattern even though the task was too difficult for school age children. For future study, grouping cue or auditory cue can be provided to raise the salience of non-adjacent elements. Thereby, it facilitated the acquisition of non-adjacent dependency. For instance, the first and final element can be chosen from similar octaves while selecting the intervening element from other octaves. This resembles the research done by Creel et al. (2004) and Gebhart et al. (2009), which demonstrated the effectiveness of perceptual similarity on learning of non-adjacent dependency.

In the experiment, about one fifth (17.9%) of the participants traced the pattern according to the melodies, which concerned about the ascending or fluctuating contours, of tone strings. This may suggest that identifying melodies was less demanding than memorizing the exact musical tones. Future study can test the non-adjacent dependency rule with musically meaningful sequences, such as melodies, as in Endress (2010). The study examined the learning of two non-adjacent major triads or minor triads intervening by musical notes, the ornaments. Since the participants needed not to recognize and memorize the exact musical tones, it showed that the learning non-adjacent dependency of melodies was much easier.

Distance of non-adjacent dependency

Newport and Aslin (2004) suggested that acquisition of non-adjacent dependency interacted with the elements and distance of structures. For instance, computation of non-adjacent dependency began in segments, such as consonants and vowels, rather than syllables. The increase in distance reduced the strength of acquisition of non-adjacent dependency. In our study, the two-element middle X tone sequences separated the non-adjacent structures to a longer extent than that in Gómez’s study (2002). This raised the difficulty on acquisition of non-adjacent dependency. However, it is unlikely that participants would learn the non-adjacent dependency even the non-adjacent elements are intervened by one middle element. It is because we cannot reduce the distance of musical tones as short as a segment unit like speech stimuli can. Moreover, research on non-adjacent dependency of one musical tone distance also generated negative result (Creel et al., 2004).

Hearing ability of participants

Since the experiment was carried out in normal school days, there were problems of background noises. The noise would distract the participants' attention from focusing the musical tones played. In future study, the experiment can be carried out in sound booths or sound attenuated room to avoid the background noise. It will also limit the distractions in the environment. Even though we carried out hearing screening, it was invalid because background noise exceeded 41.2dB SPL. Only 1000Hz, 2000Hz and 4000Hz were screened but 500Hz at 25 dB HL could not be screened reliably (ANSI, 1991). The frequency of musical notes from middle C to high C is from 261.6Hz to 523.3Hz. Even if the participants passed the hearing screening, it is possible for them to suffer from low frequency hearing loss, which impaired their perception of testing tones. To deal with it, we have adjusted the loudness level of the speaker so that the participants were able to listen to all musical strings at the beginning of the training. Although the hearing screening could not screen out participants with low frequency loss, the hearing ability would not affect their performance in the musical tone experiment.

There is possibility that some subjects may suffer from amusia, which is musical tone deafness (Ayotte, Peretz & Hyde, 2002). However, data regarding this was not collected. Amusia would badly impair the participant's musical tone perception, which is the key factor for the task. Yet, the prevalence of amusia is low (5% in United State, Hyde & Peretz, 2004) and it is unlikely to have effect on the results.

Total exposure duration to musical tones

More exposure time suggested better learning of non-adjacent dependency (Newport & Aslin 2004). The training phase that enabled participants to learn non-adjacent dependency usually lasted for 18 to 20 minutes (Creel et al., 2004; Gómez, 2002; Newport & Aslin, 2004; Onnis et al., 2003). Gebhart and colleagues (2009) even exposed the participants to a 60-minute training session. The 10-minute training in the current study may be insufficient for the participants to acquire the non-adjacent dependency. However, overexposure to musical tones would lead to boredom or decreased sensitivity of tones. Further studies are needed to investigate the optimal duration of training for school-aged children.

Conclusion

The result could not lead to two conclusions, which are there is no rule learning of non-adjacent dependency in Chinese population and the non-adjacent dependency is domain specific to linguistic stimuli only because of the following reasons. The learning of non-adjacent dependency is difficult. Individuals would not shift their focus from adjacent dependency to non-adjacent dependency when they could not find out any rules pattern embedded. The low variability of intervening elements and use of nonspeech stimuli even put more constraints on learning of non-adjacent dependency. The participants were required to recognize the exact musical notes and distinguish them against other musical tones in the testing. Without an ability of absolute pitch, it was hard for the participants to learn the non-adjacent dependency of musical tones in the design. These issues should be addressed for the study of non-adjacent dependency of nonspeech stimuli in Chinese population.

Acknowledgement

I would like to thank you my supervisors, Dr. Anita Wong and Dr. Linda Iao, for their supportive guidance. I would also want to show my gratitude to the participating primary schools and organization. They are, presented in alphabetical order, CCC Kei Wa Primary School, SKH St. John's Primary School, Wong Tai Sin Government Primary School and Yuen Long Kam Kwong Church. I am also grateful to my family and friends for their unlimited support.

References

American National Standards Institute. (1991). *Maximum permissible ambient noise levels for audiometric test rooms.* ANSI S3.1-1991 [Revision of ANSI S3. 1-1977 (R1986)]. New York: Acoustical Society of America.

Aslin, R.N., Saffran, J.R., Newport, E.L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321-324.

Ayotte, J., Peretz, I., & Hyde, L. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain, 125,* 238-251.

Benguerel, A.-P., & Estdal, C. (1991). Absolute pitch and the perception of sequential musical intervals. *Music Perception, 9,* 105-120.

Chomsky, N. A. (1965). *Aspects of the theory of syntax.* Cambridge: MIT Press.

Creel, S.C., Newport, E.L., & Aslin, R.N. (2004). Distant melodies: Statistical learning of non adjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 1119-1130.

Ding, J. (2008). On the courses of the grammaticalization about Chinese temporal adverb "cengjing". *Research in Theoretical Linguistics, 2*(1), 15.

Endress, A.D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica, 135,* 182-190.

Gebhart, A.L., Newport, E.L., & Aslin, R.N. (2009). Statistical learning of adjacent and non adjacent dependencies among nonlinguistic stimuli. *Psychonomic Bulletin & Review, 1* (3), 486-490.

Gómez, R.L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431-436.

Gómez, R.L., & Gerken L. (1999). Artificial grammar learning by 1-year-olds leads specific and abstract knowledge. *Cognition, 70*, 109–135.

Gómez, R.L., & Gerken L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*(5), 178-186.

Gómez, R. L., & Maye, J. (2005). The developmental trajectory of non-adjacent dependency learning. *Infancy, 7*(2), 183–206.

Grunow, H., Spaulding, T.J., Gómez, R.L., Plante, E. (2006). The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders, 39*(2), 158-170.

Hsu, H.J., Bishop, D-V.M. (2011). Grammatical difficulties in children with specific language impairment: Is learning deficient? *Human Development, 53*(5), 264-277.

Hyde, K.L. & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*, 356-360.

Johnson, S.P., Fernandes, K.J., Frank, M.C., Kirkham, N.Z, Marcus, G..F., Rabagliati, H.m & Slemmer, J.A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy, 14,* 2-18.

Kemler Nelson, D.G., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A., & Gerken, L.A. (1995). The Head turn preference procedure for testing auditory perception. *Infant Behavior and Development, 18*, 111-116.

Liu H.T., Zhao, Y.Y., Li, W.W. (2009). Chinese syntactic and typological properties based on dependency syntactic treebanks. *Poznań Studies in Contemporary Linguistics, 45*(4), 495–509.

Marcus, G.F. (2000). Pabiku and Ga Ti Ga: Two mechanisms infants use to learn about the world. *Current Directions in Psychological Science, 9* (5), 145-157.

Marcus, G.F., Fernandes, K.J., & Johnson, S.P. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*(5), 387-391.

Marcus, G.F., Vijayan, S., Bandi Rao, S., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77-80.

Newport, E.L. & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48,* 127-162.

Onnis, L., Christiansen, M.H., Chater, N. & Gómez, R. (2003). Reduction of uncertainty in human sequential learning: Evidence from Artificial Grammar Learning. *Proceedings of the 25th Annual Conference of theCognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates, 887-891.

Onnis, L., Monaghan, P., Christiansen, M.H. & Chater, N. (2005). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in non-adjacent dependencies. *Proceedings of the 26th Annual Converence of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates, 1045-1052.

T'sou, B., Lee, T., Tung, P., Man, Y., Chan, A., & To, C.K.S. (2006). *Hong Kong Cantonese Oral Language Assessment Scale.* Hong Kong: City University of Hong Kong.

Saffran, J.R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science, 6*(1), 35-43.

Saffran, J.R., Aslin, R.N., &Newport, E.L. (1996). Statistical learning by 8 month olds infants. *Science, 274*, 1926-1928.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*, 27-52.

Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language, 35*, 606-621.

Saffran, J.R., Pollak, S.D., Seibel, R.L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*, 669-680.

Schellenberg, E.G., & Trehub, S.E. (2003). Good pitch memory is widespread. Psychological Science, *14*(3), 262-266.

Ward, W.D. (1999). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music*, 265-298.

San Diego, CA: Academic Press.

Xia, Q. (2011). About the Grammatical Meaning of “Yijing”. *Journal of Ningbo University (Liberal Arts Edition)*, *24*(2), 32-35.