THE ICONICITY OF CONSONANTS IN ACTION WORDS

by

XINJIA PENG

A THESIS

Presented to the Department of East Asian Languages and Literatures
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Master of Arts

June 2013

THESIS APPROVAL PAGE

Student: Xinjia Peng

Title: The Iconicity of Consonants in Action Words

This thesis has been accepted and approved in partial fulfillment of the requirements for the Master of Arts degree in the Department of East Asian Languages and Literatures by:

Zhuo Jing-Schmidt	Chairperson
Kaori Idemaru	Member

and

Kimberly Andrews Espy	Vice President for Research and Innovation; Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2013

THESIS ABSTRACT

Xinjia Peng

Master of Arts

The Department of East Asian Languages and Literatures

June 2013

Title: The Iconicity of Consonants in Action Words

Saurssure argues that the relationship between form and meaning in language is arbitrary, but sound symbolism theory argues that there are forms in language that can develop non-arbitrary association with meanings. This thesis proposes that there is a sound symbolic association between consonants and action words. To be more specific, a stop sound is likely to be associated with the action of percussion and a continuant sound with continuing movements. Evidence for such an association was found through three empirical studies. The findings of two experiments revealed that such an association is motivated by the gestures when pronouncing the consonants and by their phonetic features. A study of the verbs in Teochew dialect also revealed a similar sound symbolic association existing in the colloquial language. This thesis was conducted to direct attention to the use of empirical methods to investigate sound symbolism in real language.

CURRICULUM VITAE

NAME OF AUTHOR: Xinjia Peng

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Beijing Foreign Studies University, Beijing, China

DEGREES AWARDED:

Master of Arts, 2013, University of Oregon
Bachelor of Arts, English, 2009, Beijing Foreign Studies University

AREAS OF SPECIAL INTEREST:

Chinese Linguistics
Teaching Chinese as a Foreign Language

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, University of Oregon, Eugene, Oregon, 2011-2013

GRANTS, AWARDS, AND HONORS:

Graduate Teaching Fellow, EALL, 2011-2013

Siegel Graduate Travel Endowment Fund Scholarship, EALL, 2012

Second-class Scholarship, Beijing Foreign Studies University, 2008

First-class Scholarship, Beijing Foreign Studies University, 2006

ACKNOWLEDGMENTS

Without my advisor Professor Zhuo Jing-Schmidt's support, this paper would not have been possible. I am most grateful for her for giving me the wonderful suggestions on this project and other supports in all ways imaginable.

I am also most thankful for Professor Kaori Idemaru for being on my committee and giving me suggestions on the revision. I have also received some very constructive criticism from the cohorts in the Linguistics Department, who were kind enough to let me present on the CogLing Workshop. I have also benefited from the help of cohorts in the Linguistics and Psychology department, Paul and Job.

I also need to thank the generous scholarship offered by the Siegel Travel Fund, which has made my travel back to China possible. And I also feel fortunate to have the family I have in China, who went through troubles to help me gather linguistics materials of Teochew. I also want to thank all the people that have participated in my research, and my friends who helped referring them to me. My friends in Eugene were also the sunshine that helped me get through the raining seasons during my writing process. Thank you, Jennifer, Hsinchin, and Lydia!

Last but not the least, my husband Zhen and my parents, for being so loving and caring unconditionally.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I
INTRODUCTION

'The sign is arbitrary.' The famous quote from the founder of modern linguistics Saussure (1916) appeals to us instantly, because as language users, we accept it as a fact that trees are called *trees* and dogs are called *dogs* probably by pure coincidence. If you happen to speak more than one language, you are also likely to know that the same concept is expressed in distinct words in different languages. For the concept of dog, it would be *dog* (English), *perro* (Spanish), *inu* (Japanese), *gou* (Chinese), *kringmerk* (Inuit) in different languages. And this provides further evidence for the arbitrariness of language. However, as you contemplate on this issue of the relationship between the signs and the signified, you notice that non-arbitrary association exists with onomatopoeias, which are words that can vividly describe sounds. The non-arbitrary mapping between onomatopoeias and their meanings is so patent, that even if they appear in a language you do not speak, their meanings can mostly be guessed when placed in a context. Besides the case of onomatopoeias, in a lot of languages, there are also recurring elements in words that seem to be associated with certain meanings. For instance, in English, there is a root-like element 'gl-,' which appears in words relating to the concept of 'light,' such as 'glitter,' 'glisten,' 'glow,' 'glare' etc.

Through a series of experiments starting from the 1930s, linguists and researchers have been trying to refute the Saussurean claim of language arbitrariness by collecting actual language data and people's responses in a multitude of perceptual experiments. Sapir (1929) identified a highly symbolic meaning contrast between words containing the sound *i* and *a* in the semantic domain of size; Köhler (1929) reported a sound-shape correspondence experiment, in which when asked to match two words 'maluma' and 'takete' to a curvy shape and spiky shape, most people would match 'maluma' with the curvy shape and 'takete' with the spiky shape.

In response to the Saussurean belief of language arbitrariness, Bolinger published *The Sign is Not Arbitrary* (1949), in which he proposed there is some kind of 'innate connection between form and meaning' in language. The iconicity of signs in language, though not omnipotent, still surfaces as counter-evidence of language arbitrariness to some extent. More recently referred to as sound symbolism, the descriptive work on the

iconicity of language has reached a wide range of languages, and more so in the past two decades. A quite thorough bibliography by Akita (2012) identifies published work on sound symbolism found in a large number of languages, and these languages belong to more than twenty different language families – the ones that are most cited among these works include those of Austronesian languages (McCune, 1985; Blust, 1988), English (Magnus, 1999), and Japanese (Hamano, 1998). As for experiments on sound symbolism, researchers have demonstrated, through experiments, that there is a cross-influence between vowels and size of object (eg. Koriat & Levy, 1977; Tarte & Barritt, 1971; Taylor, 1963; Vetter & Tennant, 1967). The discussion of consonants has rarely been brought up until quite recently, which is concerned with their influence on the perception of shape (Westbury, 2005; Aveyard, 2012).

In my recent investigation towards the reduplication forms of my hometown dialect – the Teochew dialect, I came across some iconic expressions describing actions. I thus hypothesized that there might be some consonant-action associations within this dialect. I am thus curious in what way those consonants are associated with certain meanings. But forming a hypothesis from scant language data from one language can be risky; I thus felt a need to investigate the issue with an empirical approach. In this study, to contribute to the understanding of consonant symbolism, I will conduct research to gather empirical data to look at the sound symbolic association between consonant (in word) and action.

The process of this investigation can also be seen as a reaction to the experimental tradition of studies on sound symbolism, which has been a segregation of the two approaches: the semasiological perspectives and the onomasiological perspectives of the association between meaning and form. By starting from establishing an association from sound to meaning (semasiological), I will show that the same association is also accessible from the other direction – meaning to sound (onomasiological), which would hopefully increase the validity of my argument sound symbolism exist with real language. In Chapter II, I will review the most representative literatures on both how previous experiments have approached the study of sound symbolism, and the findings revealed in work concerning various languages; I will also introduce two theories: the gestural theory and the acoustic theory, which are theories that are frequently cited to

explain the sound symbolic association in language. In Chapter III, after presenting criticism towards previous methods, I will give a brief sketch of the experimental methods that I formulated for the purpose of this study. In the following three chapters, three empirical studies that I conducted will be illustrated in details to support my hypothesis about the sound symbolic association between consonant and action. Two experiments will show how the proposed relationship between consonants and actions is identifiable in sound-meaning mapping and in a meaning-sound mapping; a study on a verb list of a Chinese dialect also shows that in real language such a sound symbolic relationship exists. Chapter VII will conclude the study, by reviewing the findings from the three empirical studies, and some discussion on their contribution towards theories on sound symbolism.

CHAPTER II
LITERATURE REVIEW

2.1. Defining the Unit in the Study of Sound Symbolism

In the research of the phenomenon of sound symbolism, the discussion concerns how meanings are associated with one phoneme (a consonant or a vowel) or a cluster of several phonemes (such as gl- in English). Such a unit is probably first termed by Householder (1946) as a *phonestheme*, which refers to 'a phoneme or a cluster of phonemes shared by a group of words which …[have] some element of meaning or function [in common], though the words may be etymologically unrelated' (in Abelin, 1999, p.4). Jakobson and Waugh (1979) also acknowledged the existence of such units as being more 'expressive than cognitive' in nature. If there are truly meanings associated with these phonesthemes, it would challenge the idea of morphemes as the minimal meaningful units in a language (in Blust, 1988, p.54) - which is why among other names, McCune (1985) points out that 'submorphemes' would not be appropriate. Most of the studies also refrain themselves from going into the debate of whether phonesthemes are smaller meaningful units when compared to morphemes.

However, in my opinion, the two terms should not be treated as conflicting concepts, because they diverge in a lot of aspects. First, the meaning of a morpheme in a language is set, because it is assigned by language users; the meaning of a phonestheme in a language is generated by a collection of words that share the same phonesthemes. As a result, morphemes are employed purposefully when people create new words; as for phonesthemes, I am somewhat skeptical about their productivity. Second, morphemes are ubiquitous, but the occurrences of phonesthemes are sporadic, cross-linguistically speaking. Phonesthemes, if they do exist, can be considered as special type morphemes, relevant especially in the discussion of sound symbolism.

But all in all, due to the controversial nature of phonesthemes, in the literature review, when discussing the meaning units that are sound symbolic, I will try to use the original terms used by each author. In my own investigation, I will refer to these units simply as phonemes before they are shown to be phonesthemes.

2.2. Previous Experiments on Sound Symbolism

The experimental tradition of sound symbolism studies was started by Sapir (1929) and Köhler (1929). Although both of the studies can hardly be called scientific by modern standards, they have brought to our attention that the iconicity of sound exists beyond the scope of onomatopoeias. Sapir's (1929) experiment showed that when contrasting the sound /i/ and /a/, people tended to associate /i/ with small object and /a/ with big object, suggesting that there is a sound-meaning association between vowel and size. In his published work *Gestalt Psychology*, Köhler made a similar remark, that 'the names of things and events which are visually or tactually perceived, have originated on the basis of [the] resemblances [of the tactual chracteristics of a sound and the perceived tactual characteristics of a thing].' Köhler especially believed that such a phenomenon is most strongly observed in 'primitive languages'(p.224), by which he might refer to languages that seem to only serve the functions of basic everyday communication in a more primitive culture. He then described a simple experiment, in which he asked participants to match two sounds 'baluma'[1] and 'takete' to two objects (in Figure 1) - a curvy object (on the left) and a jagged object (on the right). In this experiment, as the readers may have predicted, participants made the 'correct matching' beyond chance level, which is to match the 'curvy sounding' word baluma to the curvy object, and the 'jagged sounding' word 'takete' to the jagged object. Such a result showed that besides the overarching picture of language arbitrariness depicted by Saussure, there do exist

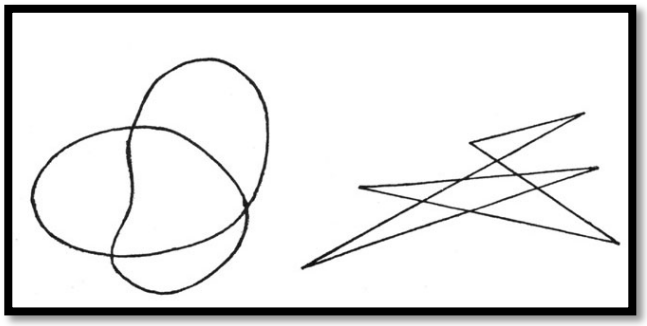

Figure 1. Curvy and jagged objects in Kohler's experiment (Köhler, 1929, p.225).

[1] 'Baluma' was changed into 'maluma' in 1947 by Kohler to avoid an association between 'baluma' and the curvy object aroused by a similar English word 'baloon.'

iconic elements in language. These elements have non-arbitrary association between sound and meaning.

2.2.1. The '/i/ & /a/' Experiments

Following the idea of Sapir's (1929) experiment contrasting the meaning of /i/ and /a/, a strand of similar experiments has been conducted by later psychologist to explore this particular issue. Whereas Sapir restricted the stimulus vowels to only /i/ and /a/, and the meaning association to only SIZE, Newman's (1933) experiment incorporated five vowels *a, i, u, e, o* and six consonants *p, t, k, d, b, g*; they were tested for the degree of SIZE and BRIGHTNESS they represented. After pronouncing word pairs that differ only in one vowel, the researcher would give the participants two contrastive meanings, and see which word of the pair is a fit for which meaning, and then participants would record their choice on paper. By carrying out multiple times of comparison on a large number of participants, Newman ended up with a relatively consistent order for both vowels and consonants – from smaller to bigger, and brighter to darker, they were *i-e-a-o-u* and *t-p-k* or *d-b-g*. He proposed that the perception of the degree was based on 'three objective factors: articulatory position, acoustic frequency, and size of oral cavity - the small-to-large symbolism follows the sequence of receding tongue position, decreasing characteristic high frequency, and increasing size of the oral cavity.' He also proposed that vocalic length of the vowels is also another factor that might have induced such perception. As for consonants, Newman also discussed 'mechanical factors at work.' If arranging the consonants by their perception from small to large, they also follow a sequence of articulatory position, from labial, dental to palatal.

The success of Sapir's experiment and Newman's experiment was said to rely largely on the method they adopt. Bentley and Varon (1933) criticized that because of a force-choice method was used in both experiments to match meanings with non-linguistic words, subject did not have the freedom to choose from a variety of meanings – instead, they were only allowed to choose two contrastive degrees on a continuum of meaning. As Bentley and Varon put it, subjects might even had 'learned under forcing to apply a half-formulated rule with respect to sounds and magnitudes' (p. 83). To test such a hypothesis, they designed an experiment, in which subjects were not forced with such meanings.

Results showed that the subjects did not respond to non-words as actively and precisely as when they were forced to associate the words with different degrees of a meaning. They then conducted another experiment, which gave subjects a context to enunciate the meanings they perceive from such words. When they were informed that the words mean a certain degree of angular, large, soft, /a/ always contrasts with /i/ to represent 'rounder,' 'larger' and 'softer.' Even though not all qualities would yield universal results like the three qualities stated above, there is a good chance that there are other unmentioned qualities out there that will always be associated with /a/ on one end and /i/ on the other. Bentley and Varon commented that the way /a/ and /i/ differed from each other was by their acoustic values, just like how volume and pitch are used in music and speech to represent different qualities. On the other hand, they have admitted that 'the feel of the tongue, jaw, chest, and other parts of the body, contributes its large share to the sound as heard. As the context is favorable, this tactual and kinesthetic component, variable in many directions and through wide ranges, probably serves in quite essential ways to convey its own meaning (p.86).'

Later experiments of this strand include Vetter and Tennant's (1967) and Tarte and Barritt's (1977) – the latter one used statistical analysis (which strikes one as being more accurate), and it showed that when trying to link a sound symbolic association from sound to size, only vowels were relevant to the judgment, but not consonants as Newman suggested. The results of these two experiments were similar to the previous finds, that /i/ was associated with smallness, and *a/o* were associated otherwise. Vetter and Tennant (1967) contributed such unanimous results to the oral gesture cues emitted when producing the sound, whether the word was printed or read.

2.2.2. The 'Takete & Maluma' Experiments

The 'takete & maluma' experiment by Köhler (1929) as laid out above, has also been replicated in the following decades. Unlike the '/i/ & /a/' experiments, the interest towards which has gradually died down in the past decades, Köhler's experiment continues to intrigue psychologists and linguists with the unresolved mystery: how people were able to come to make the same association between sounds and shapes even with varied language backgrounds. Replicated experiment has been more or less consistent to the original one, and they have seen success among English speakers,

Swahili speakers and speakers of certain Bantu dialect groups (Fox, 1935; Davis, 1961; Holland & Wertheimaer, 1964; Ramachandran & Hubbard, 2001). Even though there is an exception – reported by Rogers and Ross (1975), their experiment yielded contradictory results among a group of people called the Songe living in Papua New Guinea – it was argued that their language might be biased in a different way on such sound-shape association (Roger & Ross, p.106). In a broader context, however, the association proposed by Köhler can be universally observed.

More recently, there has been rising interest in researching what aspect of the word has induced such consistent results, that 'takete' is associated with angular shape, and 'maluma' is associated with curvy shape. Among these studies, the debate over whether it is the consonant or the vowel that is sound symbolically decisive is yet to be called conclusive. Ramachandran & Hubbard (2001) speculate that it is from the shape of the mouth that participants build such connection, and thus the vowel is the main cause of such sound symbolic association. In their experiment, over 95% of the subjects matched 'kiki' with jagged object and 'bouba' with curvy object. Staring from such speculation, a similar experiment was replicated by Maurer et al. (2006). They had four pairs of items that has one curvy and one angular, to be matched with fours pairs of sounds, which had words only with round vowels and words only with unrounded vowels[2]. The reseachers argued that their results support the theory that vowels determine the sound-shape association (confirmed on both a group of adults and a group of two-year-old toddlers), but in my opinion, the design of the stimulus did not help rule out the influence of consonants. If we look at the stimuli, we notice that in each word, there were only vowels and consonants of the same feature – meaning they were tested to provoke the same type sound-shape association. Even though the results seem to conform to the hypothesis, it still failed to compare and determine the strength of sound-meaning association provoked by vowels versus consonants. Kovic et al. (2010)'s sophisticatedly-designed experiment is also subject to the same criticism. Regardless of the tools and instrument they adopted, there was only a pair of sound stimulus that was being tested (the words 'riff' and 'mot'

[2] The pairs used in this experiments are ' baamoo & kuhtay,' 'kaykee & boobaa,' 'gogaa & teetay,' and 'tuhkeetee' & maaboomaa.'

were selected to represent angular and curvy objects respectively). The results of this experiment, which confirmed their hypothesis, could not be overwhelming evidence that vowel shape decides the shape of the object. It only showed that there is a congruent relationship between 'riff' with angular objects and 'mot' with curvy objects; the components (/i/ and /o/ here), when occurring with other consonants, might not, however, arouse the same association.

At the same time, experiments that hypothesize consonants (Westbury, 2005; Nielsen & Rendall, 2011; Aveyard, 2012) being the key factor that influences meaning-shape association also suffered from similar problems, but different measures have been taken to minimize the influence coming from vowels. Through his implicit tasks experiment, Westbury was able to establish a relation between stop consonants (he used /b/, /p/, /k/, /t/ and /d/) and spiky shapes, and also between continuant consonants (he used /l/, /m/, /n/) and curvy shapes. Such a diversity of the consonants provides strong evidence for the categorization. Nielsen & Rendall's categorization is slightly different. They used more detailed consonant categories, which is strident consonants [k], [t], [p] versus sonorant consonants [l], [m], [n] and the experiments had the results as expected: that the strident consonants are associated with angular shape, and sonorant consonants are associated with curvy shape. Aveyard's (2012) experiments yielded similar results to Westbury's with an even larger inventory of consonants. All these experiments had taken measures to minimize the interference of the vowels. For instance, Westbury used CVC word shape for all the stimuli, in which both of the consonants were of the same feature, while the vowels were randomly assigned to meet the phonotactic rule of English. Therefore, the results calculating the association based on the feature of the consonants do not immediately equal to that of the vowels. In Nielsen and Rendall's experiment, they used CVCV words, which also had consonants of the same feature to fill in the two consonant slots, while vowels were randomly and evenly chosen from only /a/, /e/ and /u/ to avoid the biased /i/ and /o/. Aveyard used three-syllable CVCVCV words, and he filled the position of the vowels with those taken from an inventory of twelve vowels randomly, and in each word, the choice and the order of the vowels were always different. In this respect, these experiments were more successful because they ruled out the influence of vowels and focus on the feature of the consonants

Therefore, in order to investigate the relative influence of vowels and consonants on sound-shape correspondence, the factors of vowels and consonants must be varied systematically. As an alternative, research can also be set up where stimuli are composed of words that have consistent feature (both consonants and vowels are considered 'curvy' or 'angular') and inconsistent features ('curvy consonants' with 'angular vowels' or the other way around), and participants' different responses to the stimuli are recorded to compare their quantitative difference, and decide whether the change of vowels or the change of consonants have more significant effect.

Following such idea, if we compare more closely the rates of 'correct choices' (meaning those that conform to the expected sound-symbolic association) in tests which used words with vowels and consonants of the same feature (Maurer's 'kiki' and 'bouba') and those that did not (Nielsen & Rendall's experiment whose vowel inventory have only /a/, /e/ and /u/), we can notice that the first type of experiment had a accuracy rate of over 80% (for adults), while the latter was only nearing 80%. Such comparison might suggest that the combination of both vowels and consonants have the strongest effect on sound-shape correspondence, and manipulating any one factor would diminish the effect. In fact, a condition in Nielsen & Rendall's experiment created inconsistent-feature pairs from Maurer's consistent-feature pairs – the consonants were swapped between the curvy sounding word and the angular sounding word. The results showed that if accuracy was judged by if it matches the prediction by the consonants, the tests had an accuracy rate of almost 80%, which means that if accuracy is judged by the vowels, it is only slightly more than 20%. Such results might be implying that consonants have a stronger sound symbolism effect than vowels.

2.2.3. The Design of the Sound Symbolism Experiments

The methods different experiments have adopted to establish their theory of sound symbolism vary. The most prototypical experiment method is to offer two stimuli with two objects to match to, such as Köhler's 'takete & maluma' experiment. Similar methods has been adopted by Ramachandran & Hubbard (2001), Maurer et al.(2006), Nielsen & Rendall (2011), which are often criticized for being too obvious in their research objective. Other studies sought to develop a framework that is subtler. Some experiments, such as Westbury (2005), were designed basing on the theory that

'congruency facilitates recognition.' His experiment asked participants to identify non-words from a mixture of real words and non-words, which were presented in the middle of a frame, which is shown as either having jagged edges or curvy edges. The assumption was that since these non-words have similar syllable structure, theoretically, they should be recognized within approximately the same reaction time, unless there are interfering factors present. Results showed that when non-words were presented with a congruent picture frame (such as when a 'jagged sounding' word was paired with a jagged frame, or a 'curvy sounding' word with a curvy frame), they were recognized in a faster reaction time than when they were presented in an incongruent frame. Therefore, one can either say that congruency facilitates recognition, or that incongruence impedes recognition. Westbury's experiment triggered participants sound symbolism mechanism without consciously evoking it.

Imai et al. (2008), Kovic et al. (2010), Kantartzis et al. (2011) and Aveyard (2012)'s respective experiments, on the other hand, exploit on the idea that 'congruency facilitates learning.' In the different learning tasks in the studies mentioned above, when participants were presented with study materials that had confirmed sound symbolic value, after several rounds of training, they always learnt these material more accurately, or made the accurate matching more quickly when compared to with material that were deliberately swapped to be incongruent.

2.2.4. Experiment Based on Actual Languages

The earliest experiments that investigate sound symbolism of real languages often draw on antonyms of a language and test how well the non-native speaking participants can match meaning correctly to the items given (Brown & Nuttall, 1959; Kunihira, 1971). The prototypical antonym experiments, starting from 1934 as introduced in Brown et.al (1955), have more recently shifted into experiments that look at how would the sound symbolic matching condition facilitate learning (Nyggard, 2009). All these experiments have tried to make a point that the characteristics attached to these antonyms that make them distinguishable, memorable are inherent in the language investigated. Brown & Nuttall even suggested that if 'antonyms evolve toward phonetic contrasts appropriate to their semantic contrast, languages in general 'evolve toward a state of richer phonetic symbolism, since words whose sounds are 'fitting' would be more easily retained and

hence aided to survive (1959, p. 444).' Though not every thing above can be correct, the last statement about the preservation due to sound symbolism is quite reasonable. It is thus also reasonable to say that sound symbolic words in a language might retain the way they are for a very long time. They even might have been created sound symbolically in the first place.

An experiment that is most discussed in sound symbolism is probably by Berlin (1994), which used actual names of birds and fish in Huambisa (a language spoken in Peru) for subjects, who are American university students to match them to either a fish or a bird - one of the animal was the actually the animal that takes the name in the language). The subjects made the correct match beyond chance level, as predicted.

Experiments of this sort have their significance. While experiments discussed throughout 2.2.1 to 2.2.2 used only non-word as stimuli, the validity of their argument is only valid within non-word. But with evidence shown in this section, one can argue that sound symbolism not only exist with people's perception towards sounds. This kind of perception is used productively in word creation too. The next section will look at the descriptive work on the sound symbolism in actual languages as evidence against the claim that sound symbolism are observed only with non-words.

2.3. Studies on Sound Symbolism in Languages

2.3.1. Swedish and English

Looking at words that start with 'gl- ' in English, one is instantly stricken with how the majority of them are related to the meaning of 'light.' This is the most intuitive piece of evidence for the term phonesthemes, which is introduced by Bolinger (1950). Bolinger proposes that the significance of phonesthemes is best tested by the lexical frequency of the investigated item in the target language. Abelin (1999) adopted such a methodology and she gathered evidence that among 8300 morphemes of Swedish, 1250 words (root morphemes) were, by her definition, sound-symbolically motivated. One feature of these words is that they contain a consonant cluster which is a phonestheme – in another word, there are more than two or three roots in the language using this phonestheme that express a common meaning. For the 37 initial consonant clusters found in Swedish, 36 of them were phonethemes.

Abelin further put these consonant clusters into order by the total number of motivated roots attached to the phonesthemes (p.80), the proportion of motivated roots to all roots of a certain phonesthemes (p.87), and the types of meanings that were popularly represented (p.90), such as 'pejorative (found with pj-, sl- and sn- words, etc.),' 'sound (found with fn-, skv-, kn-, gn- words, etc.),' 'long thin form (found with spj-, str-, spr- words, etc.),' 'quick or strong movement (such as fl-, spr-, vr-, sr- words, etc.)' and 'wetness (such as skv-, spr-, sm- words, etc).' These categories are in fact the more common categories that are likely to be sound symbolic motivated. Sapir (1929) has discovered that generally, regardless of language affiliation, words that are highly intense are more likely to be richer in phonetic symbolism (p.225) – thus in this case, the pejoratives. For other categories, they are either easily motivated by onomatopoeias, or they correlate with the popular sound-shape association speculated and investigated by psychologists and linguists as reviewed in the last section. In Abelin's experiment, she also gave native speakers production task and recognition task of neologism made up by these phonesthemes. The results supported the productivity of the Swedish phonesthemes, with some being more outstanding than others.

2.3.2. Austronesian Languages

When studying Austronesian languages, the oceanic linguist Robert Blust (1988) discovered reoccurring patterns of certain components of words having association with certain meanings. For instance, Table 1 a list of possibly related words from different Austronesian languages, which all share the root –pak

Table 1. Words sharing the root '-pak' in different Austronesian languages.

Language	Word	Meaning
Bontok	dospak	'slap someone's face with the open palm of one's hand
Kankanay	tampak	'salp in the face, box the ears of'
Hokano	ripak	'sound of a door slamming, a plate breaking, etc'
Tagalog	pakpak	'wing; applause'
Tagalog	p-ag-akpak	'noise of flapping wings'
Bikol	upak	'to clap'
Malay	cepak	'sound such as that of water lapping on a beach'
Balinese	tepak	'flat of the hand; hit with the flat of the hand'

What these words share in common is they all end in 'pak,' which is not a word when alone, nor is it a root in any of the language. But when cross-comparing these

languages, one will find ‘pak’ behaves like a root meaning ‘slap, slam’ in these Austronesian languages. Besides ‘pak,’ dozens of other roots have also been identified by Blust in these languages (Blust, 1988, Appendix 1).

The existence of such roots is not an imaginary hypothesis. The fact that they have been productive in forming words in different languages, and that they are passed on from generation to generation and remain relatively stable is strong evidence that they are important components within the Austronesian language family. The fact that native speakers of different Austronesian languages can become aware of the semantic grouping induced by the roots (experiment described in p. 51-52) also aids in recognizing the psychological reality of these roots.

Blust calls these roots sound symbolic roots, but not every root is as onomatopoeic as ‘pak.’ According to Blust, ‘sound symbolism’ should not be viewed as narrow as only onomatopoeias. He argues that for the Austronesian roots, the ‘sound-meaning relationship is defined by the analytic operation of recurrent association’ (p.65); these roots should be considered representations of ‘sound symbolism’ as well.

In the roots Blust discovered, there seems to exhibit well-established patterns of sound-meaning association. For consonants, their meanings are significantly influenced by the position of the segment. For instance, ‘initial voiced stops generally signal a louder or deeper sound in onomatopoetic roots than the homorganic voiceless stop (p.44)); ‘final stops symbolize a sound with abrupt termination, whereas the homorganic nasals symbolize a gradually diminishing resonance (p.45),’ which contrast is universally found in a lot of languages, including Japanese (see section 3.3). There are also cases that are less attested in Austroesian languages, such as the association of word-initial velar consonants to ‘sounds produced in the throat, hence to vocalization (p.45),’ but it is also not unfamiliar in English and Swedish. There are also sound meaning associations manifested in certain configurations of sound segments, which is termed by Blust as ‘gestalt symbolism,’ because symbolism is conveyed by the arrangement of certain sound into a certain shape of a word. For instance, a lot of words in Austronesian languages which are related to the meaning of ‘wrinkled’ have the skeleton of a velar stops as the initial consonant and a liquid as the second consonant.

2.3.3. Japanese

Japanese has an outstanding sound-symbolic system – the major components of which are mimetic words, or *giongo/giseigo/gitaigo* in Japanese. Mimetic words are adverbs and nominal adjectives that are used to 'symbolize manners or psychological conditions' (Hamano, 1998, p.12) but not onomatopoeias. Hamano (1998) has a systematic investigation on the Japanese sound symbolic systems highlighting the mimetic words. The following section will summarize important findings of this book that are relevant to the current study.

In Asano's (1978) dictionary of mimetic words, 1600 entries are recorded. This large inventory of mimetic words is said to compensate for the semantic under-differentiation of verbs in Japanese (Hirose, 1981 in Hamano, 1998, p.1).

There are different forms of realization of Japanese mimetics. Hamano boils them down to a one-syllable root (CV) and a two-syllable root (CVCV). After analyzing data of the mimetic words, she observes that the phonemic constituents of the mimetics function in different semantic domains and represent different semantic features, and her theory about the CV root (p.64) and CVCV root (p.104) is essentially very similar. Table 2 a chart for the constituents and their semantic representations in a CV root:

Table 2. Hamano's illustration of the different semantic representations by different types of phonemes in a CV root

Types of phonemes	Semantic representation
Initial consonant	tactile nature of the surface and the type of movement.
Voicing of obstruents	Massiveness, heaviness.
Palatalization	Uncontrolled energy and childishness
Vowels	Size and shape of movement
Diphthongs	circular movement
Vowel length	length

As this paper is concerned with mostly consonants, I will briefly summarize the main points of Hamano's arguments about the consonants in mimetics. With the initial consonants, Hamano decides that all stops are concerned with hitting in different manners, while /p/ and /b/ are both also associated with explosive movements. The continuants /h/, /s/ and /z/ are also related to a kind of movement semantically, but they are different in that the former deliver a sense of 'abruptness,' while the later are more smooth. The difference with stop consonants and continuant consonant can be better

illustrated with the following example: if one says a ninja disappears /pat-to/,' it means he disappears 'all of a sudden'; but if he disappears /sat-to/ it means he disappears 'hurriedly.' A detailed discussion of the semantics of initial consonants can be found for the section on CV roots (p. 86-99), and CVCV roots (p. 133-174).

Other than the initial consonants, consonants also have an effect if they are presented as the final elements of the mimetic words. CVCV roots can end with a /-ri/ ending, or a nasal, a glottal stop – the latter two are also acceptable ending with the CV roots. Hamano states that there are also relatively stable semantics attached to each of these final elements. For instance, the nasal ending indicates 'the action involves elastic objects or is accompanied by a reverberation'; the glottal stop ending indicates 'the movement is carried out forcefully or vigorously in a single direction'; /-ri/ implies 'quiet ending of the movement' (p.106).

Hamano points out (which is probably apparent at this point to the reader) that such association is iconically motivated, and how certain semantics come to be associated with certain words are determined by the way they are pronounced. For instance, Hamano claim, within the stop consonant group, 'the semantic split into the tactile nature and the movement seems to be based on two phonetic properties,' which are the place of articulation and manner of articulation (p.169); with CVCV roots, the first vowel deals with the initial shape, while the second vowel deals with the changed shape after movement, which is chronologically linear.

2.3.4. Asiatic Language(s)

Recorded by Diffloth (1979), there is a category of words that are highly iconic in many Mon-Khmer languages, express meaning iconically such as *thon-thon* in the following sentence (p.51).

(1) Khasis: 'ŋii ʔeec ʈhoŋ ʈhoŋ la ka rii'
English translation: 'We love our country dearly (ʈhoŋ ʈhoŋ)'

They had been called phonaesthetic words, or descriptive words, but Diffloth decides that they shall follow the Jakobson tradition and be termed 'expressives' (p.49). In terms of the part of speech, they modify verbs and are thus adverbs; but they also have other features. For instance, most of them manifest in reduplication form, or other morphological patterns; they also have peculiar phonological configuration that are not

common in regular words, which make them easy to identify. In the same article, Diffloth also touches upon the iconic mechanism employed in expressives found in these languages. Take the expressive *gp-gweep* describing 'movements of bat wings' for example, Diffloth comments that a medial 'w' conveys the sensation of free movement; and as for the final consonant - which is to tell us how the sensation ends, 'the final stop' in this example 'suggest a neat, abrupt ending, and the bilabial closure suggests wide or heavy contact,' representing the way the sensation ends' (p.56).

Diffloth, however, believes that expressives are not a sort of 'pre-linguistic' form of speech, but they are at the end a sort of 'post-linguistic' stage where 'the structural elements necessary for prosaic language are deliberately re-arranged and exploited for their iconic properties, and used for aesthetic communication. (p. 58)'

2.3.5. Chinese Languages

Iconic expressions in Sinitic languages did not draw much attention, especially outside of Chinese-speaking academia until very recently. Research interests in this branch also seem to follow a different path of investigation from those studies done in other languages, focusing more on the morphology and phonology, as Diffloth's study. Such a difference might be caused by the absence of recognition of how these expressions are sound symbolic, until Mok's (2001) dissertation titled *Chinese Sound Symbolism* came along. Mok subsumes onomatopoeias and 'state words' – traditionally viewed as two distinct lexical classes within the Chinese academia – under the concept of sound symbolic vocabulary. Even though Mok's study focuses heavily on the phonology of sound symbolic vocabulary, on how onomatopoeias and state words both 'exhibit certain phonological 'abonormalities' which set them apart from the regular lexicon (p.13), her investigation, together with other previous studies on Chinese reduplication, which include both onomatopoeias and state words, provides evidence for the argument of the similar sound symbolic nature of both classes of words. In Meng's (2012) study, she likewise argues that both should be considered 'ideophonic words,' which are lexicalized items to manifest the expressive mode of language use (p.20).

Some examples of Chinese onomatopoeias and state words are given below for the reader to gain a better picture of the manifestation of Chinese sound symbolic words.

(2) Onomatopoeias (from Mok, 2001):
/ti ta/ ‘tick-tock; sound of dropping water’ (Mandarin);
/tsʰiŋ tsʰiŋ tsʰam/ ‘whisper’ (Cantonese);
/pi li pia lia/ ‘sound of a small explosion due to over-dryness’ (Hakka).

(3) State words (from my dialect informants):
/tɕiau swan tɕiau swan/ ‘sour-ish’ (Dalian);
/poʔ poʔ tɕʰoi/ ‘crispy’; (Cantonese)
/ɕiaŋ pʰən pʰən/ ‘fragrant’; (Mandarin)
/xuaŋ-bu-la-tɕi/, ‘yellow (in an annoying way)’ (Tianjin);

Probably due to the heterogeneity of the Chinese languages, the diverse inventories of phonemes have prevented the study of sound symbolic vocabulary in Chinese to be approached at a more fine-grained level, such as phonestheme. The common morphological and phonological characteristics of these vocabularies across various Chinese languages have been explored relatively thoroughly, but few attempts have been made to investigate how such symbolism is achieved. This might be an area worthy of future attention.

2.3.6. Summary

Sound symbolism in language can be found in but not restricted to the languages above. Akita (2012) has an exhaustive bibliography on sound symbolism in different languages from more than twenty language families (by the categorization of Ethnologue). From the examples given above, we can see two diversifications of sound symbolic representation. On the one hand, there are languages like Swedish, English and the Austronesian languages that find accumulating evidence of phonesthemes; on the other hand, expressions that mimic the mental perception of events and objects (expressives or ideophones) are found in Japanese, Vietnamese, and Chinese. In Japanese, interestingly, phonesthemes have been found in expressives (see 3.3 in this chapter). It is thus reasonable to speculate that sound symbolism in the languages that have expressive might be exhibited on an even more fine-grained level – phonesthemes. Departing from such a hypothesis, this paper will examine the case of possible phonesthemes exisiting in a Chinese dialect, which by Kalgren’s account, is one of the most archaic dialect of the Sinitic languages.

2.4. Explanations for the Sound Symbolic Association

After looking at the recurring pattern of sound symbolism observed in make-up words and sound symbolic expressions revealed to use in language, such as Japanese

(2.3.3) and Vietnamese (2.3.4), we notice that the extent to which various (clusters of) phonemes are similarly utilized to convey meanings across languages is astonishing. The universality of sound symbolism might thus go beyond the observation of the universal occurrence of sound symbolism in different languages; it could be translated to the thesis that there is a universal mechanism behind the perception and production of sound symbolic words. The accounts given in different studies point to two possible theories behind sound symbolism: an acoustic theory of language perception, and a gestural/motor theory of language perception.

For instance, in Newman's (1930) interpretation of the research results of his experiment trying to establish an association between vowel, consonants and some semantic categories, he proposed that for vowels, the articulatory position, acoustic frequency, the size of oral cavity and the vocalic length of the vowels might be what the perception is based on; and the different degree perceived by consonants also 'follow a sequence of articulatory position (p.68).' In Bentley and Varon's (1933) experiment examining the association between vowel and size, they speculated that the acoustic values of the vowels might influence the perception; but at the same time, 'the feel of the tongue, jaw, chest and other parts of the body' also 'contributes to' how the sound is perceived (p.86).

The description given above represent the two different theories proposed. Back then, researchers used to seek a co-effect explanation to account for their findings, but more and more recently, they tend to use only one of either theory. In the following the two theories will be discussed.

2.4.1. The Acoustic Theory of Language Perception

This theory tries to explain some of the reoccurring sound symbolic association observed in previous studies from the angle of phonetic or acoustic information provided by the sound stimuli. Developing his argument from previous scholars, probably also including Tarte (1982), who suggested that formant frequencies of vowels may be critical in the phonetic symbolism, Ohala (1984) captures the sound symbolic association between sound and various meaning categories relating to size in his frequency code theory, which argues that 'an innately specified frequency code' within human beings allow us to make association between high acoustic frequency with 'small vocalizer' and

between low acoustic frequency and 'large vocalizer'. From an ethological point of view, he argues that F_0 of voice 'indirectly convey an impression of the size of the signaler,' because F_0 is related to the size of the vibrating membrane. From what Morton (1977) observed, the primary meaning of SMALL-BIG is also associated with other secondary meaning in communication, such as 'small vocalizer' as associating with secondary meaning of subordination, good will, and the 'large vocalizer' with aggression and dominance. After reviewing some similar uses of F_0 cross-language and cross-cultural, Ohala proposed that frequency code is innate for mammals.

Applying Ohala's frequency code theory to previous studies, it conforms to the finding, not only for vowels, but also for consonants[3], such as Newman's (1933). However, the meaning categories related to the theory are essentially restricted to meanings that are associated with the meaning category of size. Besides the affective meanings Morton mentions, Tsur (2006) also elaborates on other related notion (such as LOW-HIGH, THIN-THICK, LIGHT-HEAVY) and physical dimensions (such as SLOW-FAST, NARROW-WIDE), which, in his word, are 'analogous and co-varying' (pp.910-1). To use frequency code to explain other sound-meaning associations would thus pose a problem. The gestural theory of language perception might thus be an alternative.

2.4.2. The Gestural Theory of Language Perception

In the earliest experiments like Sapir's (1929), 'gestures' and other related terms have been mentioned repeatedly to account for the sound symbolic association, such as 'vocal gestures (Sapir, 1929, p.225), 'visual/tactual resemblance' (Kohler, 1929, p.224), 'cues from articulation' (Vetter & Tennant, 1967) and etc. In the investigation of the Polynesian languages in the 1860s, J. Rae also identified the languages as having a gestural nature. He commented that there are identifiable connection between the sign and the thing signified. To be more exact, 'the pronunciation of each separate syllable indicates a certain configuration of the organs, and that a particular configuration has positive analogies, direct or indirect, with the actions or objects indicated' (paraphrased

[3] High frequency consonants are: voiceless obstruents when compared to voiced ones, ejective higher than plan stop, dental, alveolar, palatal and front velars when compared to labials and back velars (Ohala, 1984, p.9).

in Paget, 1930, pp.157-8). One example he used is the word *mimi*, which can find counterparts in Sanskrit, Greek and Latin which also means 'the voiding of urine,' and he established an analogy between the compressed lips in pronouncing the word mimicking the action of squeezing. Rae associated the use of gesture with the languages being the 'most primitive form of human speech.' He commented that 'the primitive articulation and significant sounds only expressed force, form and movement' (in Paget) before other meanings are 'ingrafted.' Even though there was no further elaboration of this idea from Rae, we can speculate by his rationale, the earliest form of language must have originated from gestures.

Rae's description of such a phenomenon might be most appropriately captured by Wilhelm Wundt's gesture theory of language in 1904 (see Paget), which proposed that human language was an elaborate development from basic expressive movement (summarized in Kendon, 2004, p58). More evidence supporting the gestural theory of language origin surfed up in Paget's (1930) work and Jóhannesson's (1949). By describing the gestural connection between sound and meaning in roots existing in various Indo-European languages, their works presented evidence of the use of pantomime/gestures in word creation. Paget compared some 130 common Polynesian roots found by Rivet, with roots of various other languages, such as Indo-European, Semitic, Sumerian and archaic Chinese, and he found an impressive number of them sharing the common meanings. And as such 'coincidence' could not be conveniently dismissed as borrowing, the gestural theory might be more convincing. Exactly because of the existence of the same mechanism, and people produced sounds which imitate the postures and gestures of the action with the organs of articulation (Paget, p.174), we can now still find constructed roots that are shared among so many distinct languages.

In fact, philosophers in as early as the 18th century have already put forward such an idea. Among them are de Condillac (1971/1756) and Vico (1953/1744). After nearly 300 years, the gestural theory is still gaining new proponents. In a recent article, Gentilucci & Corballis (2006) discuss how human speech should have evolved from gestures rather than vocals. For one thing, 'nonhuman primates have little if any cortical control over vocalization, which is critical to speech' (p.951); for another, citing Talmy (in press), if language had evolved from a vocal system, vocal features would be more

likely to be utilized to represent according action, such as 'rising pitch to indicate climbing motion, rapid speech to indicate fast motion,' which is not a practice commonly observed in known languages. They thus argue that our language system evolved with the visuo-manual system (p. 951).

Neuroscientists have proposed the existence of a mirror system in human brains, which enables us to match actions that we observe to the execution of such action. The theoretical existence of the mirror system thus provides a physiological basis for the imitation of gestures to the action that people want to describe. However, such information is still insufficient to account for the transfer of gestures from hands to mouth. However, Rizzolatti et al., (1988) have discovered that when monkeys do a grasping action with hand or mouth, or when they are observing such an action, neuron firing can be observed in area F5, which thus establishes some connection between the hand and mouth. In Gentilucci et al. (2001)'s study, they suggested that these neurons might be functionally involved in preparing the mouth to grasp the object when the hand grasps it, and Gentilucci and Corballis (2006) also believe that they may have been instrumental in the transfer of the gestural communication system from the hand to the mouth.

The same process is recently termed by Gallese as 'embodiment,' which focuses on the 'mapping of conceptual knowledge within our sensory-motor system' (Gallese & Lakoff, 2005, p.2). In Gallese & Lakoff's (2005) framework, the activation of stimuli (perception) will cause neurons to fire in the F5 area – a pre-motor area that controls the movements aiming at object, and they discuss how such stimuli are exclusive referring to action where a target is involved (p.9). Similarly, Gentilucci & Corballis (2006) provide initial evidence of such connection in the context of ingestive movement. They speculate that gestures of the mouth might have been added to the manual system to form a combined manuofacial gestural system.

2.5. Summary of Literature Review

Sound symbolism, which might first strike people as something random and irrational, is indeed more grounded than speculation. Through experiments that investigate certain sound-meaning associations, recurring patterns surfaced from the various studies. Together with discoveries of sound symbolic elements in different

languages, the body of studies on sound symbolism is trying to make a point that sound symbolism is something that can be investigated and attested in empirical studies. And this paper will thus contribute to the body of research by conducting empirical studies to testify my hypothesis on the sound symbolic association between consonants and action word.

While one takes inspiration from the design of previous experiments, one also needs to review them critically. In 2.2.3, I discussed the framework used by previous studies. The more recent invention of sophisticated experiment designs (which aim at being subtle with their research goals) are adopted in response to the criticism given to the matching task design – which was the most commonly used technique in studies in the previous decades. They were under attack for (1) forcing the options onto the participants, and (2) having noticeably contrastive stimuli and thus disclose the goal of the experiments. However, I believe complexity is not the only solution to achieve subtlety of research goal. In the next chapter, I will discuss methods to be used in this study. The research methods I used for each of the experiments all seek to be simple in their design, while also try to be implicit with the research goal.

Besides looking at findings from previous research, we also briefly look at the acoustic theory and gestural theory of language perception, which were developed to account for the various sound-meaning associations observed. I will expect these two theories will also help to enlighten my research finding in this study.

CHAPTER III
RESEARCH OVERVIEW

3.1. Overview of Research Methods

I hypothesized that there is a sound symbolic association between consonant and action. However, at the beginning of the study, it was not entirely clear what kind of associations exist between them. To avoid generating subjective speculation, more reliable associations would be drawn from data gathered from empirical research. The data, to be specific, were to be people's perceptions towards consonants in a sound symbolic context in relation to action. Unlike previous studies, which often restricted the boundary of possible meanings by giving participants the meaning category or contrastive meaning pairs to choose from, this study (in Chapter 4) wanted to give participants enough freedom to develop their perception of consonant in words when used to describe action. The stimuli of the experiment would thus have to be put into a sound symbolic construction.

In the Teochew dialect, the '…/kio/(叫)' construction is such a construction (see 3.2). Unlike its Mandarin counterpart, the '… 叫' construction, which is an onomatopoeic construction, the Teochew '…/kio/' construction can be used with both onomatopoeias and state words. It was thus expected that when the sound stimuli were placed in such a construction, the perception towards the stimuli would be sound symbolic. Looking at examples of non-onomatopoeic sound symbolic expressions in example (6), the construction also appears to be describing mostly action scene, which satisfies the requirement of the stimuli construction.

The responses from the participants about their perception towards these expressions were gatherer, and then analyzed for the feature they shared. Results (in 4.x) would show that in this experiment, the stop sounds were perceived to represent a percussion-like action, and that continuant sounds represent the action of continuing movement. Based on such a hypothesis, the onomasiological approach was adopted in the second study (in Chapter 5), for subjects to elicit sound from meaning – technically speaking, production of sound with condition. While the former experiments have been attacked for presenting materials that are of obvious contrastive features, usually in the form of static pictures, this study used videos as meaning stimuli. Besides presenting

authentic videos depicting the two target actions, I also included twice as much distracter videos to make the contrast less apparent. I asked the subjects to watch the video and choose from the two sounds (a stop and a continuant) – supposedly two verbs adapted from an East Asian language – to decide which one is the correct word for the action. The reason I posed them as a foreign language was because obvious non-words in previous experiments were criticized for direction more attention to the contrastive feature. I assumed that since the subjects from different language background do not have prior knowledge of this 'language', they would have to rely on their native language, or the characteristics of the words to make associations, which was why I encouraged them to make association 'based on intuition.'

The process of this investigation can also be seen as a reaction to the experimental tradition of studies on sound symbolism, which has been a segregation of the two approaches: the semasiological perspectives and the onomasiological perspectives of the association between meaning and form. By starting from establishing an association from sound to meaning (semasiological), I will show that the same association is also accessible from the other direction – meaning to sound (onomasiological). Not only does it work both ways, it is also something that is shared regardless of language background – in another word, such association between sound and meaning is independent from one's language system and universal. This would hopefully increase the validity of my argument sound symbolism exist with real language.

Since the two studies have shown that there is a strong consonant-action association, it is possible to find evidence language. A third study (in Chapter 6), went back to the Teochew dialect, and tested a modified but related hypothesis on the inventory of verbs in that language. The verbs would be tagged for both their meaning category and their sound category, to see if there is any association. Even though such an operation could be carried out with essentially any documented spoken language, but the merit of the Teochew dialect is that it is a monosyllabic language, and each word only has one initial consonants. This would greatly simplify the process of defining the sound category of each word.

3.2. Stimuli for the Perception Experiment: The '…/kio/' Construction in Teochew

In the first study, I intended to include both kinds of sound symbolic expressions observed in Chinese languages: onomatopoeias and 'state words.' The '…/kio/[4]' construction in the Teochew dialect thus allow the subjects to generate both kinds of perceptions in the face of a series of sounds/words. I thus choose the '…/kio/' construction in the Teochew dialect as the framework to be used for the stimuli.

Both onomatopoeias and 'state words' in Teochew are invariably embedded in a construction '…/kio/.' /kio/ is cognate of '叫,' which means 'calling' or 'shouting.' The construction of *'... 叫'* is relatively common in Chinese dialects. However, their applications are narrower than in Teochew. In Mandarin, the form *'... 叫'* is used exclusively with vocal onomatopoeias, such as */wa-wa/ 叫* for human moaning, and */gu-gu/ 叫* for the sound of pigeons. In Teochew, some examples from Zhang (1982, 1979) are:

(4) /k^{h}i k^{h}a kio/ 'naïve laughter' and
/hip hop kio/ 'sound of chewing, biting'

which can also be adapted into a three-syllable C_1V_1F C_1V_2 F C_2V_2 F /kio/ form as follows with the same meanings:

(5) /k^{h}i k^{h}a la kio/ 'naïve laughter' and
/hip hop lop kio/ 'sound of chewing, biting.'

Somewhat wider than in Mandarin, in Northern Wu dialects, '… 叫' is an adverbial construction with limited productivity, usually used with a conventionalized set of words. In some variety, such as the Ningbo dialect, it only has two expressions, 慢慢叫 and 好好叫, which can be translated into the English equivalent of 'slowly' and 'be good and …' The construction in Teochew can never be as adverbial-like as in the Wu dialects. However, the 'state words,' which can also be regarded as a kind of adverb, also use this construction in Teochew. They can be sounds imitating certain actions, movements, such as the sound of raining, the sound of a plane flying by, and the sound of

[4] Teochew is a tonal language. However, as this paper does not concern itself with the prosodic information of the expressions, tones of all Teochew data would be removed out of stylistic consideration.

knocking on the doors; they are also used to describe the manner of certain actions, such as:

(6) /sop sop kio/ describes a person being efficient in action;
/bi bi kio/ describes a person, especially kid, being extremely gratified when eating.
/huãĩ huãĩ kio/ describes a person being restless.

These expressions, probably originated from onomatopoeias, can be regarded as the metaphorical extension of the onomatopoeias. Concerning the stagnant connotation aroused by the term 'state words,' I would thus like to propose that this kind of pseudo-onomatopoeic phrases are functionally more similar to *expressives*, as introduced in the last chapter and defined by Diffloth (1979). The term expressives is chosen over 'ideophone' because it aims at distinguishing these words from onomatopoeias. Expressives usually aids with describing the manner of the verb, from the subjective perspective of the speaker. Looking at data of expressives from various Min dialects, a Yue dialect (Cantonese), Wu dialects (from Shanghai, Ningbo, Shaoxing), Mandarin (Beijing, Dalian, Chongqing), there are several characteristics common to expressives in Chinese dialects: (1) expressives have collocation constraints with the verb/adjective they modify. The expressive morphemes (which is usually the reduplicated part) are bound with only certain words; (2) expressive morphemes across these dialects exist only in colloquial language, and they do not have orthographic representation; (3) expressives are common to be found in the domain of sense - gustation, olfaction, audition, vision and tactition across Chinese dialects, such as:

(7) Taste: /tɕiau swan tɕiau swan/ 'sour-ish' (Dalian);
Smell: /ɕiaŋ pʰən pʰən/ 'fragrant'; (Mandarin)
Color: /xuaŋ bu-la-tɕi/,'yellow (in an annoying way)' (Tianjin);
Shape: /pʰaŋ ku ku/ 'fat looking, fat and full' (Shanghainess);
Texture: /poʔ poʔ tɕʰoi/ 'crispy'; (Cantonese)

These expressives all have a prominent affective valence. The third characteristic of expressives is thus an intriguing case to consider. In Ramachandran & Hubbard (2001), they proposed a cross-domain mapping in the human brain, which allows the conversion of visual input to the auditory center, then to the motor center which controls the movement of oral gestures which eventually produce sounds. By such a theory, it is highly possible that input from other senses besides vision can also be received and transformed by the auditory center. The senses create stimuli that are then converted into

language through the above-mentioned channel, and it is transformed back into the different senses following a reverse path. Because of the existence of such a pathway, the communication of senses can be effectively achieved through the use of language. In plain words, such 'vividness' of the sensory-related expressives helps them to gain recognition, and can thus exist prominently in colloquial language.

In Teochew dialect, besides expressives in the domain of senses, there is also a considerable inventory of expressives describing scenes of actions, which echos with Weiss' prediction, that phonetic symbolism would be 'more pronounced in […] action words (1966, p.576).' In Ramachandran & Hubbard's (2001) framework, the motor system controlling hand gestures and that of the speech organs neighbor each other, and thus hand motions will most easily trigger the imitative motion in speech organs. Thus if an action is related to hand motion, it would probably produce a large number of expressions that are generated with a gestural cue. However, the extent to which other actions and motions can also be captured in such a mapping system is unknown under the given framework. In Teochew, expressives describing scenes of actions can adopt the '…/kio/' construction, but variations from this construction, which uses a verb in the place of /kio/, can also be found – except that the collocations are fixed, and for the verbs, they each have their own reduplicated sound symbolic elements. In Teochew, I have identified around twenty-five expressives that are clearly related to action or motion. And the fact that these expressives can exist so idiosyncratically in the language implies that there is a salient association between the sound symbolic elements and the action, which is easily perceivable by the Teochew speakers. This can be used as evidence to further show that the association between action/motion and sound do exist in this dialect. Some examples of Teochew expressives in the domain of action/motion include:

(8) /ŋa ŋa gun/ '(of people) crowded as if water boiling'
/pʰi pʰi pue/ 'flying around, travelling around'
/laʔ laʔ so/ 'loafing around'

Previously we have talked about how the construction of '…/kio/' is employed in the Teochew sound symbolic expressions. The words above, which use verbs or adjectives in the /kio/ position, can be seen as using a paradigmatic variation of the '…/kio/' construction, since most of them can still be replaced back to /kio/ in the verb/adjective position and have the same meaning. They are thus also sound symbolic in nature.

By examining the various manifestations of the Teochew '…/kio/' construction, we can conclude that a variety of sound symbolic expressions are attainable through this construction – be it onomatopoeias, expressives for a stagnant state or expressives for an action/motion. It is therefore an ideal construction to be used in the perception experiment to elicit responses to various consonants in a sound-symbolic context. And because the construction meaning of the '…/kio/' is language-specific, the subjects providing input for the perception can only be speakers of Teochew.

3.3. Background Information of Teochew

Teochew, also known as Swatow, refers to the variety of Southern Min dialect spoken in the area which has Teochew as its east-most boundary and Swabue as its west-boundary with Swatow as the center. Historical linguists of Chinese should not be too unfamiliar with the dialect, as it was introduced by Kalgren as 'the most archaic, most peculiar of all known Chinese dialects' (1926, p.135)– a comment that he made after his fieldwork of collecting phonological data from various Chinese dialects in the early 20th century. Evidence of such a statement can also be extracted from another work of his on the comparative and historical phonology of Chinese languages, *Études sur la phonologie chinoise* (1926). A simple summary is that Teochew has preserved many phonetic features that date it back to before the first books describing the phonology of Chinese languages.

The basic syllable structure in Teochew is (C)V(C). There are 18 consonants in the consonant inventory, 17 of which can be in the initial consonant position, while the glottal stop can only be in the final consonant position. The VC part, or the rhyme, has 72 combinations, which can be a monophthong, a diphthong, a triphthong, or monophthong or diphthong that is nasalized, having a nasal ending, or a stop ending. The discussion of the inventory of Teochew can be found in Li (2010, p.115 & p.137).

CHAPTER IV
EMPIRICAL STUDY I

To investigate what kind of sound symbolic perceptions are evoked with different consonants, I conducted a survey with Teochew speakers, using a list of different sound combinations to discover whether there are underlying patterns that influence their perception.

4.1. Method

In section 3.2, I have introduced the Teochew sound symbolic construction '…/kio/'. The construction was used to with different sounds to create the stimuli. After a list of stimuli was created, I interviewed eight participants individually about their perceptions towards these words.

4.1.1. Stimuli

In order to reduce possible interference caused by existing vocabulary, I used unfamiliar words that are phonotactically familiar to the speakers. In Teochew, a syllable is made up of an optional initial consonant, vowel(s), and an optional coda (either a glottal stop or a nasal). I thus chose nine consonants to match with four rhymes to create thirty-six stimuli for the survey. The nine consonants were /p/, /p^h/, /g/, /k^h/, /h/, /l/, /ŋ/, /s/, /ts^h/ - some of which are components of the sound symbolic morphemes of the action expressives I mentioned in 3.2 (including /p^h/, /g/, /k^h/, /h/, /l/, /ŋ/, /ts^h/), others were supplemented to provide contrast between aspiration (/p/ and /p^h/), place of articulation (/h/ and /s/), and manner of articulation (/s/ and /ts^h/). This inventory of consonants thus differs in their voicing, aspiration, place of articulation (labial, alveolar, velar and glottal) and the manner of articulation (stop, fricative, affricate, liquid and nasal). Four rhymes were chosen to match with different consonants. These rhymes were: a monophthong /a/, a monophthong with a nasal coda /om/, a monophthong with a stop coda/op/, a triphthong with a glottal stop /iauʔ/. They were chosen for their familiarity to Teochew speakers in the language system.

Besides the thirty six stimuli created with the material above, another fifteen using the same consonants but with new rhymes were added to the stimuli above to randomize the stimuli. Although ideally the number of distracters should have a two-to-one ratio to the actual stimuli, it was compromised so as to ensure the survey is controlled

within a thirty-minute time frame. Due to the specific nature of the data collection process, that it was a survey depending on the voluntary participation of informants, responses are expected to be more accurate when informants are not feeling impatient with the survey.

These sounds were then put into the sound symbolic construction '…/kio/' in their reduplication form. As for tones, considering most onomatopoeic expressions and expressives have low pitches, I gave the reduplicated part of the stimuli a seventh tone (31) except for those that end with a stop, which were given a fourth tone (2). In this way, fifty-one stimuli were created for the survey. Excel randomized the order of the stimuli, and the same order was used for each informant.

4.1.2. Participants

I conducted interviews with eight informants – four males and four females from four different age groups (those in their 30s, 40s, 50s and 60s) to make sure the sound symbolic associations are not idiosyncratic to a specific age group, with the consideration that there is an overall tendency of fluency decrease of Teochew with younger generations due to the nation-wide Mandarin promotion.

4.1.3. Procedure

I posed this survey as a dialect survey, telling the informants that they would be presented with some expressions and how they understand them. I started by giving examples of sound symbolic expressions to the informant in a sentence, as in:

'(it rains) /xa xa kio/'

Then I proceeded with the first expression on the list, and asked the informant, 'what do you think the expression 'xx /kio/' describes?' In the survey, the informants generally did not have difficulty understanding their task, and they did a great job giving me their perception of the stimuli. They were told that they are allowed to skip the items that they had no clues with, but they were encouraged to try their best in responding to all the stimuli.

Informant's responses were noted down in a chart for later analysis.

4.2. Results

4.2.1. Response Rate

Eliminating the distractor stimuli, there should an expected total of 288 responses, while the observed number of responses was 195, with a 67.7% response rate. None of the consonant group or the rhyme group had a response rate lower than 50% (see Table 3), which means that in general these choices of consonants and rhymes were perceived to be understandable when used in a sound symbolic construction in Teochew.

Table 3. Number of responses for each item (expected 8) and overall response rate for each consonant group and rhyme group

	/a/	/om/	/op/	/iau/	Average response rate for this consonant group
/p/	2/8	5/8	5/8	4/8	50%
/p^h/	6/8	5/8	5/8	4/8	62.5%
/g/	4/8	7/8	7/8	3/8	65.6%
/k^h/	4/8	8/8	5/8	4/8	65.6%
/ŋ/	7/8	3/8	5/8	7/8	68.7%
/h/	8/8	7/8	5/8	5/8	78%
/l/	4/8	8/8	7/8	7/8	81%
/s/	8/8	6/8	7/8	7/8	87%
/ts^h/	4/8	5/8	4/8	3/8	50%
Average response rate for this rhyme group	65.3%	75%	69.4%	61%	

In this survey, more responses were gathered from female informants, with an 80.5% response rate, which was higher than a 54.9% response rate of the male informants. Such observation corresponds with the common belief that females are in general more sensitive to language than males.

4.2.2. The Rhyme-Meaning Association

(1) In the /a/ group, one recurring semantic category was ‘loudness’ hinting some degree of unpleasantness, such as the sound of rain, the sound of wind, the sound of

water boiling, fast-moving traffic, noise made by a crowd, human's fluent speech, or baby crying.

(2) In the /om/ group, the associated meaning was very similar to the /a/ group. It was also 'loudness,' and from the description given by the informants, it seemed that the level of 'loudness' in the /om/ group was even higher than the /a/ group. The scenes reported by the informants were mostly different types of noise, typically hitting or bumping, which made up 27 out of 48 responses. Responses from /pom/, /p^{h}om/, /gom/, /kom/ and /lom/ were the main contributors for the perceptions. Further discussion about the meaning distribution by consonant will be laid out in the following.

(3) In the /iauʔ/ group, the meaning was exactly the opposite to the /a/ and /om/ group, with a majority of responses revealing an association with 'quieter sounds,' such as the sound of insects' movement, small human movements, whispers, grunts. In the 44 valid responses, 17 belonged to this category. For the rest of the responses, 4 responses out of 4 for /piauʔ/ reported it was the sound for hot cooking oil; 5 responses out of 7 for /liauʔ/ were described as sounds made by walking in flipflops; 4 responses out of 7 for /ŋiauʔ/ were referred to as sounds related to the movement of insects. The result was so because the meaning for /piauʔ/, /liauʔ/, /ŋiauʔ/ seemed to have been very conventionalized in the Teochew onomatopoeic framework. Yet still, they helped to contribute to the common perception of this rhyme group.

(4) In the /op/ group, the responses did not put an emphasis on the volume – whether it was being loud or quiet. However, the responses in general could be interpreted as related to the type of movement that is repeated rhythmically, such as heartbeats, boiling water, hitting on nails, horses galloping. Following this definition, there were 35 responses out of 60 valid responses that fell into this category.

The three meanings: LOUDNESS, QUIET SOUNDS and RHYTHMIC MOVEMENT, and OTHER (if none of the three options could be applied) were used to code each response given by the informants. A Pearson's Chi Square test confirmed the observation: there was an association between rhymes and meanings $X^2(9)=133.7$, $p < .001$. The null hypothesis that there is no association between rhyme of sounds and the meaning was rejected. Rhyme-meaning matching did exist in this survey. Positive associations were observed between 'loudness' and the rhyme group /a/ and /om/

(p<0.0001), between ‘quiet sounds’ and /iau/ (p<0.0001) and between ‘rhythmic movements’ and /op/ (p<0.0001). There were also strong negative association between ‘loudness’ and /iauʔ/. /, between ‘quiet sounds’ and /om/, and also between ‘rhythmic movements and /a/ and /iauʔ/ (p<0.005). The distribution of meanings in each rhyme group, and make-up of rhymes in each meaning group are shown in the Figures 2 and 3.

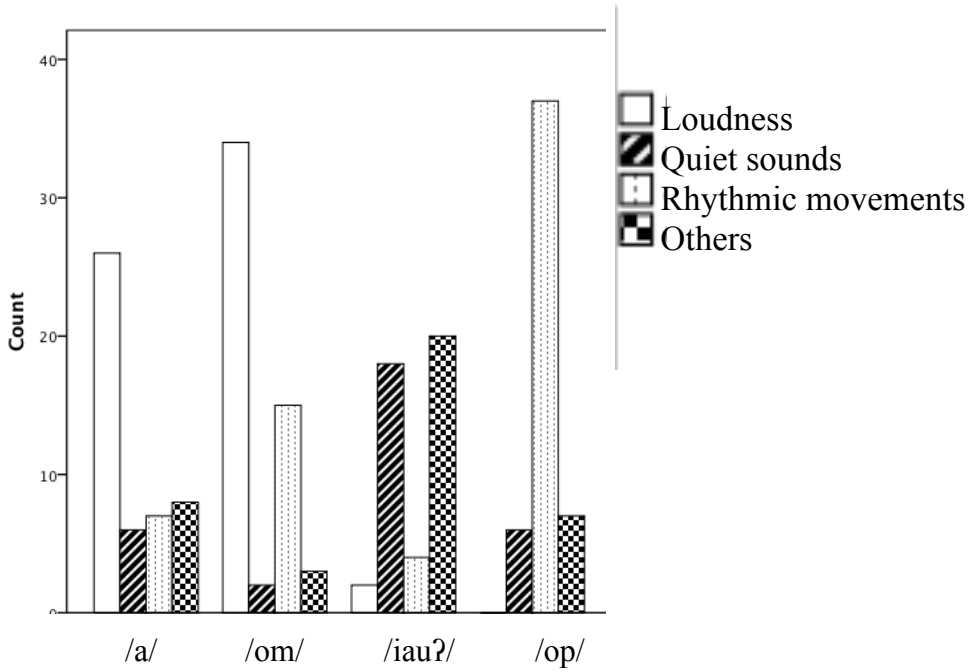

Figure 2. Different meanings represented in different rhyme groups

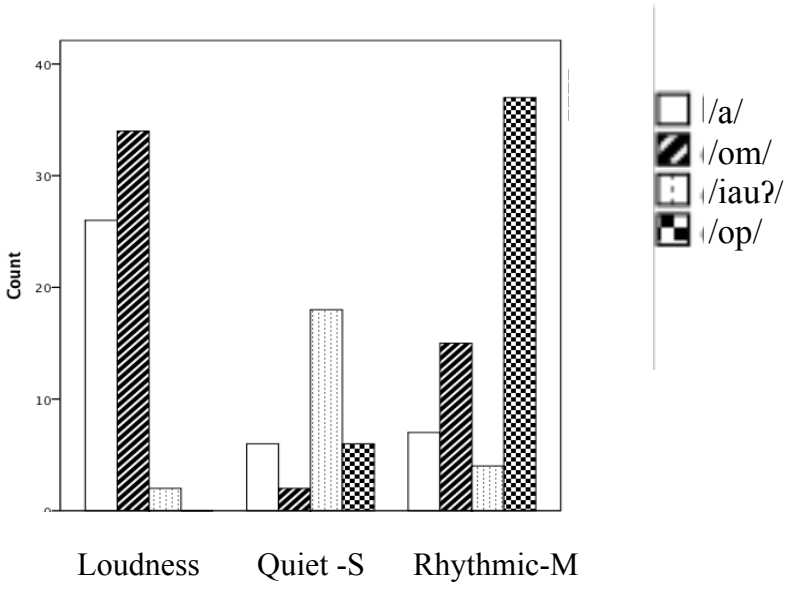

Figure 3. Make-up of rhymes in each meaning group

4.2.3. The Consonant-Meaning Association

The treatment for consonants was similar to that of the rhymes. After analyzing qualitatively the responses given by the informants, I postulated several meanings that emerged from the responses, and then each response was coded with one meaning, or 'other' if not applicable. These meanings included: VOCAL SOUNDS, PERCUSSION, ERUPTION and FAST MOVEMENT. But it was also possible that some meanings remain unidentified in my analysis.

Pearson's Chi Square test shows that the null hypothesis that there is no association between consonants and meaning was rejected, $X^2(32)=248.8$, $p < 0.001$. Positive associations were most strongly observed between the consonant /ŋ/ group and 'vocal sounds,' between /p^h/ and 'eruption,' between /k^h/ and 'percussion,' and /s/ and 'fast movement,' all of them had a p value of $p<0.0001$. Less strong but still significant positive associations could be found between /g/ and 'vocal sounds,' /p/ and 'percussion,' /h/ and 'fast movement,' $p<0.005$. Negative association could be found with /g/ & /k^h/ & /p/ and 'fast movement' ($p<0.05$), /h/ and 'percussion' ($p<0.05$), /l/ and 'vocal sounds' ($p<0.05$), /ŋ/ &/ts^h/ and 'percussion'($p<0.05$).

The distribution of meanings in each consonant group, and the occurrences of different consonants in each meaning category are shown in Figure 4 and 5.

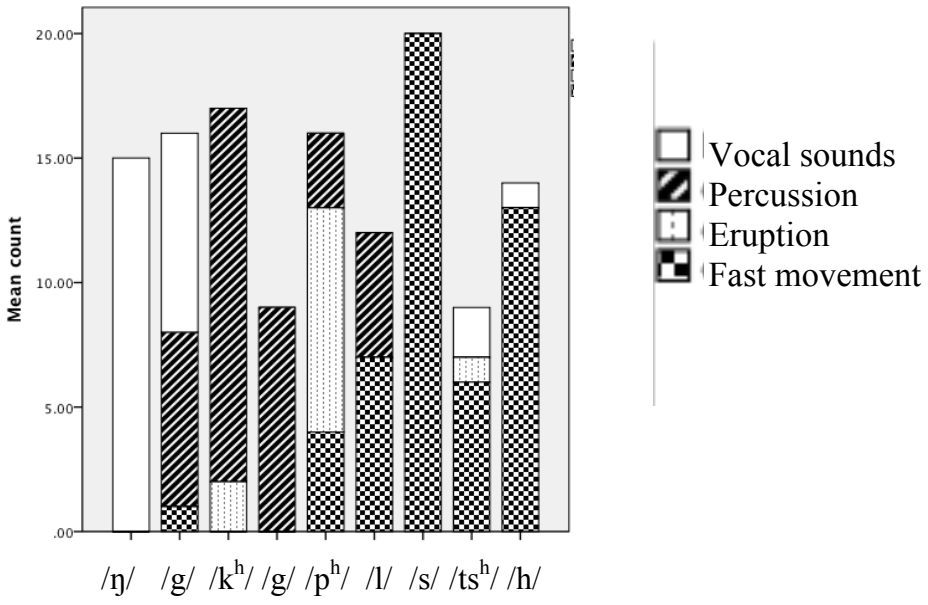

Figure 4. The make-up of meanings in each consonant group by counts

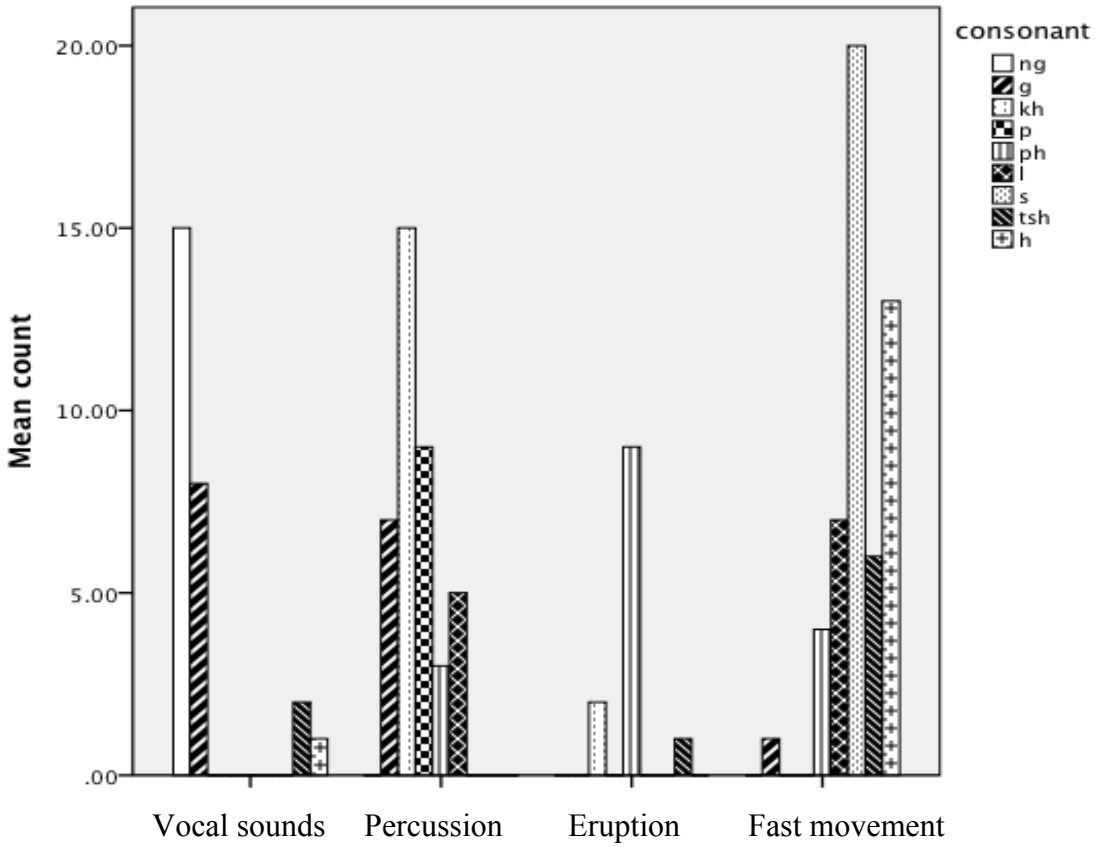

Figure 5. Occurrences of consonants in each meaning category

Of all the valid responses, (see Figure 4) 71.4% of responses in the /k^h/ group, and 56.3% of the /p/ group were related to percussion; 68.2% of the /ŋ/ group and 38.1% of the /g/ group responses were related to unclear vocal sounds; 71.4% of the /s/ responses are related to fast movement, especially of people running fast, while the /h/ group also has 52% of its responses relating to it – a huge proportion of which related to describing 'water flows,' which is 33%, which I had subsumed under the 'fast movement' category. 37% responses in the /ph/ group reported a description of 'eruption.' From Figure 4 and 5, it is clear that certain meaning associations were not exclusive to one consonant. The consonant groups that shared the same meaning also share some phonetic features. The meaning of 'percussion' could be found in the velar stops, labial stops and also the liquid /l/; for the meaning of vocal sounds, responses mostly came from velar sounds, including /ŋ/, /g/ and /k^h/; for the meaning of fast movement, responses mostly came from continuants, plus the aspirated bilabial stop. We can thus observe some connections

between the meaning of ‘percussion’ and stop sounds, between ‘fast movement’ with continuants, and also between ‘vocal sounds’ and velars.

4.3. Discussion

From the results above, we have observed non-arbitrary sound-meaning associations both with rhyme and consonant. Upon closer examination, it seems that the motivation behind rhyme-meaning association and consonant-meaning association is somewhat different.

4.3.1. The Rhyme-Meaning Association

The semantic contrast of ‘loudness’ and ‘quiet sounds’ might be explained by the phonetic difference between /a/, /om/ and /iauʔ/. The waveforms of the sound clips of the three rhymes matched with the consonant /h/ (Figure 6) shows that there is a clear contrast of sonority between /a/, /om/ and /iauʔ/. Such contrast of sonority of sounds must have been conveniently converted into the contrast of volume for the action that is being described. Phonologically speaking, the glottal stop in /iauʔ/ might have helped reduce the syllable weight of the rhyme so that even when it was a triphthong, it was pronounced not as sonorant as /a/ and /om/. And Teochew speakers might have utilized such contrast of weight/sonority to represent the volume for the actions that are being described.

Another effect of stops, as found both in /op/ and /iauʔ/, is that it reduces the duration of the syllable. The reduction of duration is reflected towards the time of actions described. And since /op/ is even shorter than /iauʔ/, such characteristics was magnified and was perceived prominently for its projection on time duration, while the reduplication assisted in forming the ‘rhythmic’ interpretation.

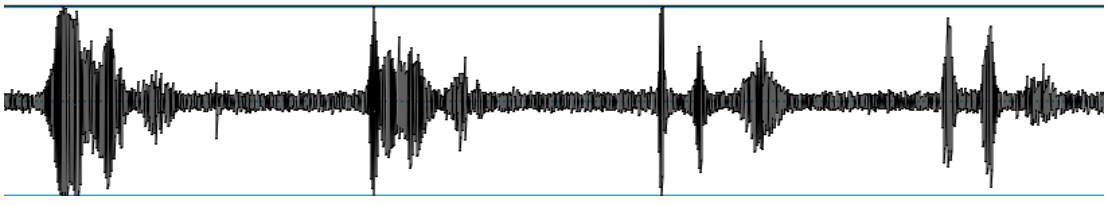

Figure 6 .Waveform of different rhymes in the /...kio/ frame

4.3.2. The Consonant-Meaning Association

In 4.2.3, the results imply that different meanings reported with these consonants might be contributed by their phonetic feature. Thus if a sound possessed more than one feature as identified above, it was likely to be associated with two meanings. And through contrasting sounds that were minimal pair of certain phonetic feature, one can have access to what meaning is associated with the feature that was being compared. Referring to Figure 4, the statistics of the /g/ group is thus interesting, because both the meaning of 'percussion' and 'vocal sound' overlapped here, while these meanings were respectively prominent in different sounds that are related to /g/ - which is /k^h/ and /ŋ/. When /g/ was devoiced and became /k/, its meaning association with 'percussion' would be strengthened; on the other hand, if its sonority was increased, the other meaning association – vocal sound, was strengthened. We may thus propose that the consonants' different phonetic characteristics are related to certain meanings in this discussion of sound symbolic words.

Similarly, the other two stops, /p/ and /p^h/ might be associated with the meaning of 'hitting and banging' due to their stop feature, with /p/ having 56.3% such responses and /p^h/ only15%. The low representation in /p^h/, similar to the case /g/ was because that the other feature of it distracted it from being faithful to the meaning. /p^h/'s aspiration feature made it comparable to other continuant sounds, whose dominating meaning implication was 'fast movement' – this meaning also received a 20% response rate from the informants. However, aspiration was not contributing solely to the meaning of 'fast movement.' If compared to other aspirated consonants in this data, which were /k^h/ and /ts^h/, /p^h/ was identified to share with them the meaning of 'eruption,' such as 'firing a canon,' 'coughing,' even though it only composed a very small portion in their responses. It is no coincidence that the afore-mentioned semantic category found in aspiration was completely absent in the unaspirated stops /p/ and /g/ and /ŋ/. We may thus propose that aspiration endows stops with the meaning of 'eruption,' and the feature of aspiration alone would contribute to the meaning of 'fast movement.'

Comparing the features with the meanings they represent, the manner in which the consonants are produced strongly resembles that of their semantic features. The scenes described above – 'percussion,' 'fast movement' and 'eruption' all correspond to

the way those representing sounds are pronounced. Stops are great imitation of the action of percussion, because in both processes, sounds are produced when the movement is obstructed. Continuants are great imitations of the action of 'fast movement,' because in fast movement strong airflow accompanies it, and the aspiration and sibilant sounds produce airflow just the same as in 'fast movement.' Across different cultures, there are expressions like ' he runs like wind.' It is thus not uncommon to associate movement with airflow. Aspirated stops are comparable to 'eruption,' because in eruption, a blockage precedes a strong airflow – just like in pronouncing an aspirated stop sound aspiration follows the stop.

The results for the experiment strongly echo what is outlined in the gestural theory of language (although it might be also appropriate to use the acoustic theory to account for the airflow). When informants are placed in a sound symbolic context for their perception of sounds, they utilize the gestural and acoustic cues of the speech organs available to them and apply these analogies to the perception of meanings. The abstracted meanings of vowels are thus related to their volume and length, while the actions of pronouncing consonants are abstracted into various meanings of actions.

CHAPTER V
EMPIRICAL STUDY II

Based on the survey results above, one contrastive feature – whether the consonant is a stop or a continuant emerges as representing distinct action types: a sound starting with a stop consonant is highly associative to a percussion action, while a sound starting with a continuant consonant is highly associative to the action of fast movement, flowing action. Such categorization has been hinted also in Hamano's (1998) study on Japanese mimetics.

If such a connection is based on the acoustic and gestural information, it is something that is universally accessible regardless of one's language background. It is thus not a coincidence that Teochew speakers would have such a similar perception towards sounds as Japanese speakers. I was thus interested in finding out whether such connection between phonetic feature and semantic representation, at least in the aspect of action, is a universal perception across speakers of different languages when tested in the other direction.

5.1. Methods

5.1.1. Stimuli

Video stimuli

Since I hypothesized that the percussion kind of action and the movement kind of action are relevant to the perception of stop sounds and continuant sound, the two types of actions are represented in the video stimuli. There were nine for the action type of percussion, and nine for the action type of movement. Each action type was further divided into three subtypes to ensure that there was a balance of different representation of the target action type. A brief description of the subtypes and the content of the videos is given in Table 4:

In the perception survey, the action type 'eruption' was identified to be the most conspicuous in the aspirated bilabial stops, which had a mixed feature of both stops and continuants. I thus decided to also include some videos on eruption to gather more information.

Table 4. Contents of videos in the perceptual experiment

Main action type	Subtype	Content of Video
Percussion	Percussion occurring on a solid surface with a large contact area	A person slams a guitar on the floor until it breaks. A hammer bangs the surface of a flat-screen cell phone. A person throws himself to a hardwood wall.
	Percussion occurring on a lax surface with a large contact area	A person plays multiple drums with hands. A person claps his hands A boy plays two basketball on both sides of his body to keep them bouncing.
	Percussion occurring on a solid surface with small contact area	A person knocks on a door A person knocks on a wooden instrument with a stick. A person hammers a nail into a concrete wall.
Movement	Fast Movement following a path	Athletes run on tracks. A dog runs through a forest. A plane flies over the sky above observer's head.
	Flow of water/speech	A person does tongue twister Rain falls from the sky continuously. Water runs in a creek.
	Outward flow (of air/gas) in an undulating motion	Wind blows over some grasses and makes a wave. Some gas coming continuously from a hole on the ground. A snake sends out his tongue, withdraws and repeats.
Eruption		A canon fires. A firework explodes in the sky. Firecracker explodes all over the place. A volcano explodes and smoke comes out from the crater.

Sound stimuli

In each video, there was a pair of sounds to be compared: a stop, and a continuant. Again, I included a total of ten stops and continuants that differ in their place of articulation, voicing, and aspiration. However, these differences are not treated as variables, but only as a way to increase the variety of stops and continuants included. These consonants were then combined with the vowel /a/ because it is probably the most common vowel across languages; and also, from previous literature, /a/ also implies 'largeness' (see 2.2.1), which is consistent with the exaggerated action manifested in the videos in this experiment. The ten sounds /ba/, /p^ha/, /da/, /t^ha/, /ga/, /k^ha/, /za/, /sa/, /ha/, /la/ were produced by the program MBROIA from the English inventory. The vowel /a/

had a 649ms duration with an even pitch of 110 Hz, but the duration of the consonants varied because they were re-synthesized by natural speech.

The interface

Each video was edited to be short (3s to 5s) and of similar pixel (640*400) on the screen. The videos were also muted to minimize possible influence of sounds produced in the video, which would lead the subjects to assume that this was an experiment about onomatopoeias. Another 44 videos were added as distractors for the experiment at a 2:1 ratio, which have all been edited by the above standards. Because there were three different pairs of stops and also three different (pairs of) continuants, they were mixed and formed into nine different combinations of 'stop v.s continuant' pairs. Each video in the main action type is matched with one of the nine combinations (the eruption action type is more random, however), and then the feature of voicing was applied to the contrastive pairs wherever applicable, and 'voicing and voiceless' alternate with neighboring pairs to increase the diversity of the sounds. In each video, the order of the stop and the continuant was decided by flipping a coin, so that not all the videos had the stop sounds as the first sound of the pair. The order of these videos with the sound pairs embedded were then randomized in Excel, and then put into a Powerpoint slide show. On each slide, there was one video, and it would play once clicked; under the video, there were two icons of speakers representing the two sound, which also require clicking to play. In the banner, an instruction was placed to remind the subjects their task. The interface of the slide show is shown in Figure 7.

5.1.2. Participants

To achieve the best result for this experiment, the researcher targeted people with minimum or no second language or foreign language. Due to limitation of actual resources, I recruited eight speakers whose native language is not English, and who had been staying in the US for less than two years. Most of the subjects also did not seem to demonstrate an exceptionally high level of oral proficiency in English. The language background of these subjects was respectively: Wenzhou dialect of China (a Wu dialect) as native language, Tianjin dialect of China (a Mandarin dialect), Japanese, Korean,

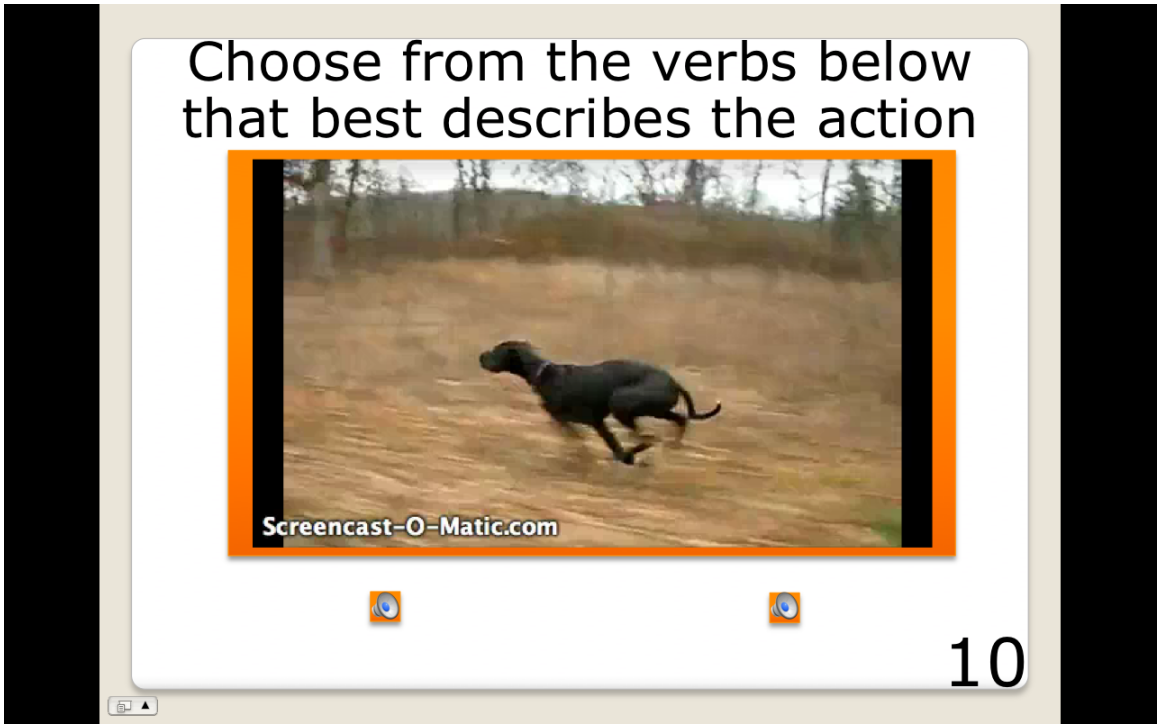

Figure 7. Interface of the perceptual experiment

Vietnamese, Arabic, Portuguese and Russian. Therefore, for each language group I only had one speaker.

5.1.3. Procedure

Before the experiment began, I informed each subject that the task for her/him in this experiment was to identify a verb from the certain Southeast Asian language that they thought is a better match for the action based on their instinct. After I demonstrated the procedure, they tried with two sample questions to show that they understand the instruction correctly. They would then proceed with the 66 videos and mark their choices on an answer sheet I gave them.

To further test whether such mapping can be universal beyond a stop/continuant feature distinction, I switched the stop-continuant pairs into a group of three sounds of the same stop/continuant feature group on the original 66 videos. For each percussion video, there are three stop sounds to be compared; for each movement video, there were three continuant sounds to be compared; two of the eruption video had the stop sounds and two had the continuant sounds. The other distractor videos were assigned randomly whether they had a continuant group or a stop group, but the occurrence of stop sounds

and continuant sounds are the same. The matching of sounds to video and the order of the sounds in each video followed the same procedure to be randomized. However, the order of the video was the same in this second experiment. Therefore, for a certain video, the subjects might be asked to choose between /la/ and /ga/ in the first experiment, the second set of videos asked them to choose one from /la/, /sa/ and /ha/ - sounds that have the same stop features (and same voicing), but differ in their place of articulation. This experiment followed the first experiment immediately and therefore was completed by the same subjects.

5.2. Results

On average subjects spent 30 minutes to finish both sets, which means that they spent an average of ten seconds to decide which sound is a better match for each question – the length seems to be of appropriate amount for the subjects to contemplate on each question.

A Pearson's Chi Square shows that there is a significance association between sound and action type matching: $X^2(2)=14.6$, $p \leq 0.001$. The continuant sounds seemed to be especially attracted to the movement action type ($p \leq 0.0001$), and the movement type repelled the stop sounds (Figure 8). On the other hand, the percussion action videos

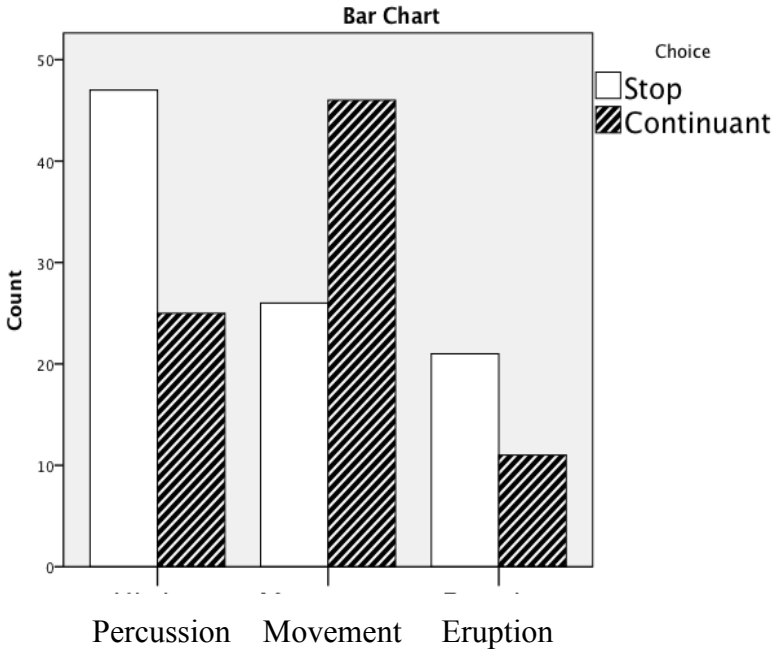

Figure 8. Consonant choices in each action type.

seemed to attract the stop sounds significantly ($p \leq 0.001$), and repelled the continuant sounds. And these findings are consistent with the prediction generated from the perceptual survey in the last section. As for the eruption action type, no significant association was observed. This was probably because in the last survey, we discovered that it was represented most conspicuously by aspirated stops, especially /p^h/. And since the pairs in our stimuli did not highlight the contrastive feature, it was difficult to equal either the stops or the continuants to an aspirated stop. The distribution of different sounds in each action type can be

The association between sound and action subtype matching was also statistically significant: $X^2(6)=16.9$, $p \leq 0.01$. However, in examining the individual correlation, only the 'fast movement' action subtype and the continuants reveals a significant correlation ($p < 0.05$).

Between choice and questions, there was also a significant association: $X^2(21)=36,3$, $p < 0.05$. The most influential decision seemed to have come from subjects' decisions on question 15 (water flow) ($p<0.005$) and 20 (gas flowing out from a whole on the ground) ($p<0.05$), which favored continuant ad stop respectively in these two questions.

I was also initially concerned about biases towards either stop sounds or continuants due to the influence of language background, but it did not seems to be the case in this experiment. A Pearson's Chi Square test showed there was no significant association between subject and their choices for the questions - $X^2(7)=11.6$, $p \geq .05$, except for the Portuguese speaker, who seemed to be extremely fond of choosing stop sounds ($P <0.005$).

For the second experiment, I didn't expect there to be significant association between the choice of a specific consonant in a certain subtype or a certain question. There was no significant association observed $X^2(4)=7.1$, $p \geq .05$ between their choice of stop sounds and percussion, except for strong attraction between 'percussion occurring on a solid surface with a large contact area' and bilabial stop ($p<0.05$), and strong repulsion against bilabial stop with 'percussion occurring on a solid surface with a small contact area' (or knocking)' ($p<0.05$). On the specific questions level, there was also no overall significant correlation with choice, except for between 'questions 2 (A hammer bangs the

surface of a flat-screen cellphone) and bilabial stop' ($p<0.05$), and also 'questions 9 (A person hammers a nail into a concrete wall) and alveolar stop ($p<0.05$).

Neither was there observed association between the choice of continuant and the movement subtype $X^2(4)=7.1$, $p>.05$, except there was a significant negative association between the liquid /l/ and the 'gas flowing out' subtype ($p<0.05$). In terms of association between specific question and choice, there was also no correlation in general, $X^2(16)=20.0$, $p>0.05$. There was however, a negative correlation between question 14 (tongue twister) and the choice of the alveolar fricatives 's/z' ($p<0.05$).

As for the subtype of eruption, since it had proven to be statistically insignificant in the first round, the results in the second round of experiment are not examined.

5.3. Discussion

The results of the experiments strongly points to the fact that when giving the task to associate sounds with actions, there is a strong tendency to associate stops with the action of percussion, and continuants, especially fricative (since the liquid /l/ did not have significant results in both studies), with the action of movement – such behavior was at least observed with the action types that had been included in this experiment.

The association between specific consonant with a certain subtype of action, or a certain scene of action, was not significant. However, some preferences could also be observed. In the subtypes, the subjects liked to associate banging with /b/ or /p/, but they did not like the association of 'knocking' with these two consonants; neither was the association between 'gas leaking' and the /l/ popular among the subjects. On the consonant and question association level, the subjects liked to associate in particular, 'slamming a guitar to the ground' with bilabial stops, 'knocking on a wooden instrument' with /t/ and /d/, but they did not like to associate tongue twister with /s/ and /z/.

In the post-experiment interviews, participants of the experiments were asked about the strategies they had used to finish the task. A majority of them were aware of the contrasting action types and the contrasting consonant types even with the presence of distracters. Many subjects reported that they established a connection between sounds that were more 'strong, intense' with actions that were 'intense, vigorous,' and sounds that were more 'soft, calm' with scenes of similar nature. Therefore, even without

linguistics knowledge, they were able to make a distinction of contrasting phonetic feature, and used it as a way to establish a contrasting relationship between videos.

Most of the subjects also reported recalling their native language, either the verb, or the onomatopoeias. And if the sound used in their language was not available in the choices, they reported choosing a 'similar sounding' word - and according to the examples given in the interview, they were choosing the sound that has the same stop feature with the one in their native language. And in making such decision, without drawing on other phonetic features, for instance place of articulation, they picked the stop/continuant feature to decide which word is a better counterpart of the words in their native language. In this experiment, we can imply that the comparability between the stop/continuant feature of sounds and the action type was accessible to the subjects. In this experiment, when the chance presented itself for the subjects to make sound symbolic connection between sounds and actions, the subjects were fully capable of developing the sound symbolic association naturally without training, and accomplish the tasks with significant accuracy.

With the knowledge that the subjects were using their native language to complete the questions, although we are not aware of the extent to which the subjects relied on their knowledge of their native language to do the tests, we might be able to hypothesize that the decisions they made about what sound is matching to a certain type of action, could to some extent, at least in terms of stop/continuant feature, reflect the choices of sounds in verbs in their respective language, which again, touches on the topic of the role of iconicity in language creation and language development, which leads me to the investigation of the next study.

CHAPTER VI
EMPIRICAL STUDY III

From the results of the first survey, I noticed in various instances in the subjects' responses, the reported scenes for the expressives coincide with the verbs having the same sound of the reduplicated elements. For instance, /liao liao kio/ was reported by five participants to remind them of people walking in flip-flops, especially when they are dragging them. A verb that has the same sound in Teochew dialect has the meaning of wearing flip-flop. It thus made me begin to wonder if some verbs in Teochew, as the case of /liao/, might bear a sound symbolic origin. If Teochew speakers were able to reinterpret the meaning of a certain word in a sound symbolic framework, then part of the sound was contributing meanings to the word. In the previous two studies, I have discussed how stop and fricative feature of consonants were associated with different types of actions; I thus hypothesize that such association also exist with the action verb in Teochew. And through this study, I would like to present evidence that the creation of action verb in this language was sound symbolically motivated.

6.1. The Imitative Nature of Sound in Words

In the previous chapters, two experiments using the semasiological approach and the onomasiological have shown that there is a sound symbolic association between consonants and action, and the phonetic features found with the consonants motivate these associations. The gestural origin of language states that we tend to use sounds that imitate the feature of the action, and in doing such imitation, the acoustic features were also taken into consideration. Looking closely, what the stops were imitating in the survey experiment is the scene when an object runs into a blockage, which causes the movement to cease. And applying this analogously to the speech organs, it also characterizes what is essential in producing a stop sound: the blockage of airflow in the vocal tract. The movement action type has a process resembling that of the fricatives– the vocal tract is partially blocked but not completely so that airflows can continue.

My hypothesis on how sound and action is sound symbolically associated should thus be modified to a finer association between stops and action of percussion, and fricatives and continuing movement.

However, if we want to look at how these sound symbolic associations might exhibit within a language, the two types of actions presented here are not categorical at all to include all the verbs within a language. But if we are to at least look at the sound-meaning association with words that have stops and fricative, analogous categorization should be devised.

Notice, that a key difference in the production of stop and fricative - whether there is a complete vocal tract blockage. If we apply an analogy of the vocal tract blockage to actions, it can be considered as whether there is substantial contact between the agent of the action, and the patient or recipient of the action.

Such definition is not completely identical with transitivity. However, I would borrow terms in the domain of semantics to help illustrate what is meant by such an analogy. Because the working definition of the differentiation of action types put an emphasis on whether there is physical contact of the two participants of the action, when there is only one participant in the action, it would be defined as the 'fricative' type of verb; when there are two participants in the action, but the force is inflicted not through physical contact, such as 'steaming' and 'magnetic attraction,' it would still be considered the 'fricative' type of verb; when there are two participants in the action, but there is no intended impact, and neither is there any intended change towards the physical form of the patient, then it would still be considered the 'fricative' type. All these 'fricative' type actions, which are not inflicted with intended impact, are analogous to the manner how a fricative is produced, which is the absence of abrupt obstruction in the vocal tract. All other cases of actions are considered to belong to the 'stop' action type.

6.2. Data and Methods

I collected a total of 411 action verbs from a Teochew dictionary. The dictionary, compiled by Junming Cai in 1976, follows an alphabetical order of the pronunciation of the monosyllabic words in the Teochew dictionary. Under each syllable sound, homophones are first categorized by their tone, then randomly without following a specific order. Each entry of word starts with a character if the word has a written representation; if not, a box is provided instead. The word is then explained in written Chinese for its meaning, and then several high frequency collocations words and phrases are given in some sentences that show how these words are used in context.

I adopted a set of standards in choosing the verbs from the dictionary.

First, due to the purpose of this research – that I am focusing on the iconicity of consonants, and how consonants are matched with different types of actions – only words that start with a consonant were included. Words that start with a vowel were excluded.

Second, only words that can find no Mandarin counterparts were included – these are basically words that have no written counterparts, plus words that are given a character but the meaning of which does not match the meaning expressed by the character in Mandarin-speaking context. In every Chinese dialect, there is always a huge inventory of words and expressions that are not transcribed. They are unanimously, and intimately referred to as *tuhua*, or 'earthy language,' for their earthiness – coarse but direct. They are also usually felt to be 'vivid' by native speakers, and such vividness is in fact another way of acknowledging language iconicity.

Third, for words that have the same meaning, if they are sharing the same initial consonant and final consonant, they were combined into one - or, they were only counted once, because there is a good chance that they are only dialectal variation of the same word, such as /hak/ and /hok/ both meaning 'to steam' was only recorded in one entry.

In this way, a list of 411 action verbs in Teochew was created for analysis. I translated the meaning of each verb into English using the Chinese definition provided in each entry.

As the investigation only concern fricatives and stops, words that start with either fricative consonants or stop consonants are selected, and they form a new list of 271 verbs. The verbs were then coded for whether it belongs to the fricative type or the stop type without referring to the original pronunciation of the words. They were later also coded for whether they were fricatives or stops.

6.3. Results

In the total of 271 action words included in this investigation, 201 of them had stop initial consonants, and only 70 has fricative initial consonant (in Teochew, the inventory of fricative consonants is in itself smaller than that of the stops. There are only three fricatives /s/, /z/ and /h/ in Teochew, while the number of consonants is much larger). There are also more stop-type verbs than fricative-type verbs observed in this investigation. A Pearson's Chi Square test shows that the association between the

consonant type and the action type is significant: $X^2(1)=21.9$, $p < 0.001$. The distribution of action representation in each consonant group is shown in Figure 9. Words that have a stop consonant tend to be associated with the stop-type action, and words that have a fricative consonant tend to be associated with the fricative-type action

6.4. Discussion

Through the preliminary investigation of the Teochew action verb, the hypothesis that there is an association between the gestural representation of the action word and the phonetic feature of its initial consonant was confirmed to have statistical significance. A perusal of the meaning and the formation of the word seem to also reveal an association between the sounds and the actions in terms of syllable weight and action duration. I also noticed that for action words that were not included in this study, there was always a predominant semantic category accompanying certain consonants, such as the meaning of 'cover' with the labial nasal /m/, and the meaning of 'rolling' with the liquid /l/. To test if there is a significant association between these sounds and meanings, future investigation is required. However, I expect that sound-meaning associations as fine-grained at this level must be more language specific than the stop/fricative feature association discussed earlier in this paper, and this paper will not attempt to elaborate on the

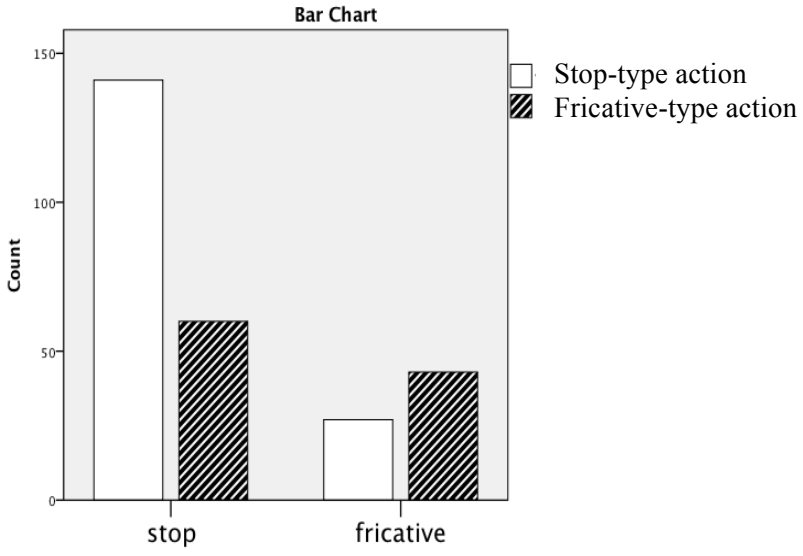

Figure 9. Distribution of meanings in each consonant group.

descriptive analysis of specific consonant iconicity observed in Teochew based on one person's intuition. However, through the investigation of real language data, I am arguing that the sound-meaning association between stop sounds and the 'stop' type action and that with fricatives – not only exists in our perception, but also in our language production - at least in one language, and possibly more.

CHAPTER VII
CONCLUSION

This research started as a quest exploring the sound-symbolic association unknown to previous research: one that investigates how consonants might be associated with action given a sound symbolic context.

The first study in this paper used a sound symbolic construction in Teochew to elicit the sound symbolic perception associated with different consonants. An analysis shows that phonetic features, instead of individual consonants motivate the sound symbolic perception. The most outstanding sound symbolic association comes from the association between stop sounds and percussion, and fricative and continuing movement.

The association was tested on people speaking other languages to examine whether such a perception is universal or language-specific. Eight subjects from eight language backgrounds showed astonishingly similar patterns of association in a perception task, in which they watched a video and decided on a sound (a stop/continuant) that better described the given action scene.

The association between consonants and action types, observed both in a semasiological study and an onomasiological study, suggests the existence of extremely prominent sound-meaning iconicity. In a sound symbolic context, when other clues are deprived of them to draw any association, people would turn to the gestural cues and phonetic information to make more 'meaningful' association.

It is thus possible that the same mechanism was used when action words were created. Weiss once speculated that phonetic symbolism might be 'more pronounced in […] action words (1966, p.576),' among other categories of words. It has now been confirmed by neuroscientists that human brains possess various mechanisms and channels to allow for different types of associations. And because we are wired in a more or less similar fashion, by the same channels, we, as human being, are able to develop similar sound-symbolic association. Among these mechanisms, the mirror system in human brain is tested to be able to transfer hand gestures to vocal gestures. Therefore, when human were able to depict certain action scenes with hand motion, it could be transferred to oral gestures, and the sound pronounced with such an oral gestures could later be used a word that denotes the action.

Which was thus why I conducted a study on such kind of association with real language: the study of action words in Teochew. The quantitative analysis of the list of Teochew verb suggests that using gestural cues and phonetic information to create words might have been an important way in which verbs in Teochew were created in the first place.

All in all, the empirical studies have provided ample evidence that there is a non-arbitrary association between consonants and actions, which is universally accessible to speakers of different languages. Although Saussure's claim of language arbitrariness strikes us as intuitive correct, I would like to emphasize, especially with my own finding through this research, that there are some aspects of language (besides onomatopoeias) that are iconic. And the iconicity of language is important, because for one thing, it can sometimes explain why some of the systematic phonological changes occur with our arbitrary language system do not affect the whole system, such as the English *cuckoo*, and the irregular phonological features observed with onomatopoeias and expressives in various language; for another, if we can identify more cases of motivated sound-meaning association with language, the origin of human language, might not be a myth as we have been made to believe.

Again, this paper sees its contribution to the larger sound symbolic literature as adding one more sound symbolic association: the association between consonants (stops and fricatives) and actions (percussion and continuing movement). The successful results yielded from this research largely rely on the different sound symbolic frameworks controlled methods adopted. Without a sound symbolic context, especially when the sound symbolic perception is in conflict with a person's knowledge of his/her own language, the results would be dramatically different. Nonetheless, through the evidence provided in the paper, I do hope that the 'arbitrariness' view towards language can be put in perspective, and that people will be more willing to accept the iconicity that exists in their language, and work together to help uncover the mythical way of how our brain works, and how language was created.

REFERENCES CITED

Abelin, Å. (1999). *Studies in sound symbolism*. PhD Dissertation, Göteborg University.

Akita, (2012) https://sites.google.com/site/akitambo/Home/biblio

Asano, T. (ed.) 1978. *Giongo Gitaigo Jien [A dictionary of mimetic words]*. Tokyo: Kadokwa.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & cognition*, *40*(1), 83-92.

Bentley, M., & Varon, E. J. (1933). An accessory study of phonetic symbolism. *American Journal of Psychology*, 45, 76–86.

Berlin, B. (1994). Evidence for pervasive synesthetic sound symbolism in ethnozoological nomenclature. In L. Hinton, J. Nichols & J. Ohala (Eds.), Sound symbolism (pp. 76–93). New York: Cambridge University Press.

Blust, Robert. (1988). *Austronesian root theory: an essay on the limits of morphology.* Amsterdam/Philadelphia: John Benjamins.

Bolinger, D. (1949). The sign is not arbitrary. *Boletín del Instituto Caro y Cuervo*, *5*, 52-62.

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *Journal of abnormal psychology*, *50*(3), 388.

Brown, R., & Nuttall, R. (1959). Method in phonetic symbolism experiments. *The Journal of Abnormal and Social Psychology*, *59*(3), 441.

Cai, Junming, (1976). *Chao yu ci dian [A dictionary of Teochew]*. Taipei: Taipei Sanmin Shuju.

Davis, R. (1961). The fitness of names to drawings: A cross‐cultural study in tanganyika. *British Journal of Psychology*, *52*(3), 259-268.

Diffloth, G. (1979). Expressive phonology and prosaic phonology in Mon-Khmer. *STMK*, 49-59.

Fox, C. W. (1935). An experimental study of naming. *The American Journal of Psychology*, *47*(4), 545-579.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, *22*(3-4), 455-479.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*(7), 949-960.

Gentilucci, M., Benuzzi, F., Gangitano, M., Grimaldi, S., 2001. Grasp with hand and mouth: a kinematic study on healthy subjects. *Journal of Neurophysiology* 86, 1685–1699.

Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, Calif.; Tokyo: CSLI Publications ; Kurosio.

Hirose, M. 1981. *Japanese and English Contrasitve Lexicography: The role of Japanese 'mimetic Adverbs.'* PhD Dissertation, University of California.

Holland, M. K., & Wertheimer, M. (1964). Some physiognomic aspects of naming, or, maluma and takete revisited. *Perceptual and Motor Skills*, *19*(1), 111-117.

Householder, F. W. (1946). On the problem of sound and meaning, an EnglisH phonestheme. *World*, *2*, 83.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54-65.

Jakobson, R., Waugh, L. (1979) *The sound shape of language*. Harvester Press, Sussex.

Jo'hannesson, A. (1949) *Origin of language: four essays*. Reykjavík, H.F. Leiftur, Blackwell. Oxford

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese Sound‐Symbolism Facilitates Word Learning in English‐Speaking Children. *Cognitive Science*,*35*(3), 575-586.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Köhler, W. (1929). *Gestalt psychology*. New York: H. Liveright.

Koriat, A., & Levy, I. (1977). The symbolic implications of vowels and of their orthographic representations in two natural languages. *Journal of Psycholinguistic Research*, 6, 93–103.

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19-28.

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior*, *10*(4), 427-429.

Li, Hongxin, (2010). *Chao yu [The Phonology of Teochew]*. Shantou: Shantou Daxue Chubanshe.

Magnus, M. (1999). *Gods of the word: Archetypes in the consonants*. Kirksville, USA: Thomas Jefferson University Press.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science*, 9, 316–322.

McCune, K.M. (1985). *The internal structure of Indonesian roots*. Jakarta: Badan Penyelenggara Seri Nusa, Universitas Katolik Indonesia Atma Jaya.

Meng, Chenxi. (2012). *A description of ideophonic words in Mandarin Chinese*. PhD Dissertation, Universiteit Leiden.

Mok, W. E. (2001). *Chinese sound symbolism: A phonological perspective.* PhD Dissertation, University of Hawaii at Manoa.

Morton, Eugene S., 1994. Sound symbolism and its role in non-human vertebrate communication. In: Hinton, Leanne, Nichols, Johanna, Ohala, John J. (Eds.), *Sound Symbolism*. Cambridge University Press, Cambridge, pp. 348–365.

Newman, S. (1933). Further experiments in phonetic symbolism. *American Journal of Psychology*, 45, 53–75.

Nielsen, A., & Rendall, D. (2011). The sound of round: evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, *65*(2), 115.

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*(1), 181-186.

Ohala, J. J. 1994. "The frequency codes underlies the sound symbolic use of voice pitch," In L. Hinton, J. Nichols, & J. J. Ohala (eds.), *Sound symbolism*. Cambridge: Cambridge University Press. 325-347.

Paget, R. (1930). *Human speech*. London: K. Paul, Trench, Trubner & co., ltd.; New York, Harcourt, Brace & company.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Rizzolatti, G., Camarda, R., Fogassi, M., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71, 491–507.

Rogers, S. K., & Ross, A. S. (1975). A cross-cultural test of the Maluma-Takete phenomenon. *Perception*, *4*(1), 105-106.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, 225–239.

Tarte, R. D., & Barritt, L. S. (1971). Phonetic symbolism in adult native speakers of English: Three studies. *Language and Speech*, 14,158–168.

Taylor, I. K. (1963). Phonetic symbolism re-examined. *Psychological Bulletin*, 60, 200–209.

Tsur, R. (2006). Size–sound symbolism revisited. *Journal of Pragmatics*, *38*(6), 905-924.

Vetter, H., & Tennant, J. (1967). Oral-gesture cues in sounds symbolism. *Perceptual and Motor Skills,* 24, 54.

Weiss, J.H. (1966), A study of the ability of English speakers to guess the meanings of non-antonym foreign words. *Journal of General Psychology*, 74, 97–106.

Westbury, C. (2005). Implicit sound symbolism in lexical access: evidence from an interference task. *Brain and language, 93*(1), 10-19.

Talmy, L., in press. Recombinance in the evolution of language. In: Cihlar, J.E., Kaiser, D., Kimbara, I., Franklin, A. (Eds.), *Proceedings of the 39th Annual Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, IL.