2010

Sound symbolism, sonority, and swearing : an affect induction perspective

Yardy, Brandon John

Lethbridge, Alta. : University of Lethbridge, Dept. of Psychology, 2010

SOUND SYMBOLISM, SONORITY, AND SWEARING:

AN AFFECT INDUCTION PERSPECTIVE

Brandon John Yardy

B.Sc. University of Western Ontario

A Thesis

Submitted to the School of Graduate Studies

of the University of Lethbridge

in Partial Fulfillment of the

Requirements for the Degree

MASTER OF SCIENCE

Department of Psychology

University of Lethbridge

LETHBRIDGE, ALBERTA, CANADA

ABSTRACT

The relationship between word form and word meaning has been debated since early Greek philosophy. Conventionally, the relationship is held to be arbitrary: that there is no natural connection between a word and what it represents (de Saussure 1959). In contrast, examples of sound symbolism undermine this linguistic tenet by demonstrating non-arbitrary word meanings conveyed in details of the acoustic signal of the words themselves. The Affect Induction model of animal communication offers a natural explanation for some forms of sound symbolism in language. According to the Affect Induction model, the physical properties of signals influence receiver affect and behavior in specific ways through relatively direct effects on core sensory, psychological and affective processes. To investigate the possible implications of this model for sound symbolism in human language, a set of studies was conducted on the classic "bouba-kiki" phenomenon. An analysis was subsequently undertaken to extend the results of experiments to several corpuses of real words classically associated with divergent affective themes. Results suggest that the Affect Induction model might account for some forms of sound symbolism, as instantiated in real word usage.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER FOUR:

LIST OF TABLES

LIST OF FIGURES

1.1 Theoretical Background

The nature of the relationship between symbols and their referents has caused vigorous debate since at least the times of ancient Greek philosophy. In Plato's dialogue *Cratylus,* Hermogenes took the position that all words were related to their referents as a result of cultural convention, whereas Cratylus took the position that words were related to their referents naturally. In defense of the position that words naturally fit with their meaning, Socrates notes that certain letters or syllables seem to naturally suit their referent by emphasizing aspects such as motion, speed and softness. Socrates was thus suggesting that simple speech sounds, or phonemes, could naturally convey meaning. Socrates does, however, go into great detail on how words can change simply due to happenstance and he does not commit fully to either the naturalist position or the conventionalist position. The controversy continues to this day.

Cratylus's naturalist position is in line with the modern study of sound symbolism (Nuckolls 1999). Sound symbolism is the idea that the symbol, or word, can inherently convey meaning via the physical characteristics of the sound itself. Onomatopoeic words such as *hiss*, *crash*, and *buzz* imitate the sounds they denote and are well-known examples of sound symbolism. However, sound symbolism runs counter to the conventional view in linguistics that words, and their referents, are not inherently related at all. Ferdinand de Saussure (1959), the founder of modern linguistics,

proposed that "The symbol is arbitrary". From this perspective, words are arbitrarily formed in local linguistic communities entirely due to happenstance; for example, in English dogs are called *dog*, in French they are called *chien* and in German they are called *hond*. If there was some inherent "dogness" that connected *dog* with the referent of a four legged canine then all languages should have converged on the same name.

Further, every word cannot be sound symbolic as there are only a limited number of speech sounds used in any given language to produce the entire lexicon; for example, the words *bet* and *get* are distinguished acoustically by only a very small difference in voice-onset time, but this heralds a large difference in word meaning (Nuckolls 1999). Although language is largely arbitrary, work in sound symbolism and related fields of study suggest that it cannot be *completely* arbitrary. Indeed, it is from some important non-arbitrary, sound-meaning relationships that we may come to understand some of the underlying processes of language change and possibly even language evolution (Ramachandran & Hubbard 2001, Reilly, Biun, Cowles & Peelle 2008).

The very act of speaking conveys information independent from the symbolic meaning of the words used. Information such as sex, country of origin, current health and emotional state are all conveyed in the physical characteristics of the speaker's sound production (Ostwald 1994). Another important inherent conveyor of information is the intonation, rhythm and stress of speech, known as prosody. Prosody can be critical for proper interpretation of otherwise ambiguous statements. For example,

2

saying "they were given directions to follow" can be interpreted differently than "they were given directions to *follow*" (Ladefoged 2001). Some tonal languages, such as Cantonese and Mandarin, use pitch to differentiate among words because they have essentially the same phonemic and written structure (Ladefoged 2001). Indeed, the number four in Cantonese and the word for death vary only in pitch; this is the purported reason why the number four is considered unlucky.

Prosody has implications for *how* words are said and interpreted; however, sound symbolism could play a deeper role at the level of *what* words, or phonemes, are used. The relationship between size and the type of vowel used is one possible example: that is, high-front vowels, such as the /ɪ/ in *bit*, involve the elevation and positioning of the tongue near the front of the mouth and are associated with smallness or being diminutive. Low-back vowels, such as the /ɑ/ in *caw*, involve the lowering and positioning of the tongue toward the back of the mouth and are associated with largeness (Sapir 1929, Ohala 1994, Ultan 1978). English examples of these size-sound associations include the words *little*, *teeny*, *tiny* as well as *huge*, *colossal* and *humongous*; though, the words *big* and *small* are obvious exceptions.

In a seminal study, Edward Sapir (1929) asked 500 participants, including seven Chinese participants, to match two objects that varied in size, with two nonsense words, that either had a /ɑ/ or a /ɪ/ vowel. For example, a picture of a large table and a picture of a small table were presented to participants and they were asked to match these images to the nonsense words *Mal* and *Mil*. Sapir (1929) found that 74.6%-96.4% of

participants preferred to match the low-back vowel word with the larger object and the high-front vowel word with the smaller object. Importantly, the Chinese participants also exhibited a 78.6% preference for matching larger objects with words that contained low-back vowels and smaller objects with words that contained high-front vowels. Sapir argued that this sound-sense relationship is fixed at an early age and probably is not related to familiarity with a language or literature because of the nonsense nature of the words and the fact that the Chinese participants also exhibited the effect.

Sapir's student Stanley Newman (1933) conducted a follow-up study that extended these findings to brightness or luminosity. Using a similar methodology to Sapir's previous work, Newman found that nonsense words with high-front vowels were associated with brightness and the low-back vowels were associated with darkness (Newman 1933). More recently, this methodology of using nonsense words has been conducted with an aim of creating more suitable brand names (Klink 2000, Klink 2003 Yorkston & Menon 2004, Lowrey & Shrum 2007). Nonsense words containing front vowels were deemed smaller, lighter, milder, thinner, softer, faster, colder, more feminine, friendlier, weaker, and prettier for any particular product than nonsense words that contained back vowels (Klink 2000). The nonsense nature of the word stimuli used in these studies prevents the referential information confound of real words; however, this control also makes these studies less externally valid and does not provide direct evidence that sound symbolism is manifest in the words of real languages.

Some anecdotal evidence on this point comes from cross-cultural examples of size symbolism where words conveying "smallness" include *teeny* in English, *chico* in Spanish, *petit* in French, *mikros* in Greek, and *shiisai* in Japanese. Conversely, words involving "largeness" include *humongous* in English, *gordo* in Spanish, *grand* in French, *makros* in Greek, and *ookii* in Japanese (Klink 2000). A more substantial analysis was conducted by Ultan (1978) on 136 languages. Ultan found that size-sound symbolism did occur for about a third of the languages surveyed; however, this finding is weakened somewhat by the fact that 65% of these languages were from Western Native languages that are known for their use of sound symbolism (Nichols 1971). Berlin (1994) found support for size symbolism in words for animals in local cultures. For example, in Huambisa, a Jivaroan language in north central Peru, words containing high-front vowels are used more often for smaller fish and birds and words containing low-back vowels are used for larger fish and birds. Further, the phonetic differences between the categories of bird and fish are salient even to naïve listeners. This was evidenced by English students who successfully matched the Huambisa word to the correct category of bird or fish at levels significantly above chance (Berlin 1994).

Brown and colleagues (1955) provide further experimental evidence that sound symbolism is found in natural languages and is not limited to nonsense words or size symbolism. Participants were asked to match an English antonym word pair with a corresponding antonym word pair from one of three foreign languages: Hindi, Mandarin and Czech. Participants matched the English antonym word pairs correctly with all three languages at levels significantly better than chance; indeed, participants were right

twice as often as they were wrong (Brown, Black & Horowitz 1955). For example, 93% of

participants matched *light* and *heavy* with the Mandarin *chi'ng* and *chung* respectively,

a result particularly consistent with Sapir's (1929) work.

Brown and colleagues (1955) argued that sound symbolism was due to universal

inter-sensory connections; in other words, sound symbolism occurs cross-culturally

because of a common pattern of cross-modal connections, such as across vision and

audition. However, Brown (1958, as cited in Nuckolls 1999) later downplayed a universal

and innate inter-sensory account of the phenomenon in favour of a hypothesis that

emphasized cross-modality learned associations based on common sense experiences

across individuals. Despite Brown's (1958) reversal of position, other studies have

replicated Brown's (1955) findings and have even continued to support an argument of

an innate inter-sensory bias shared cross-culturally (Miron 1961, Weiss 1963, Gebels

1969, Lapolla 1994). The debate is still not settled as to what extent innate and

environmental factors contribute to the universal nature of this phenomenon.

1.2   The Bouba-Kiki Phenomenon

Sound symbolism has typically been researched with an emphasis on size

symbolism; however, there are other intriguing examples of sound symbolism. One such

example is the "bouba-kiki phenomenon" initially described by the famous Gestalt

psychologist Wolfgang Köhler. Köhler identified a sound symbolic relationship between

particular nonsense words and simple line images that were either an angular self-

intersecting shape or a rounded self-intersecting shape. Köhler showed these images side-by-side to participants on the island of Tenerife and asked: "which is takete and which is baluba?" The majority of participants matched the jagged image with *takete* and the rounded image with *baluba*; the same results were found when Köhler later used *maluma* instead of *baluba* (Köhler 1947). If words are completely arbitrary, then the nonsense words used by Köhler should have been equally assigned to the two images by participants. In contrast, the subjects' consistent matching bias suggests that there is some inherent expectation of a sound-meaning relationship that encourages participants to match *takete* with the jagged image and *maluma* with the rounded image.

This inherent sound-meaning relationship also appears to be cross-linguistic and cross-cultural. For example, Davis (1961) used the same essential methodology as Köhler and found that African children (age 8 – 14) matched nonsense words with images in the same way as a control group comprised of English children; that is, angular images were matched with *takete* and rounded images were matched with *uloomu*. The African children were relatively isolated by the local geography and had minimal exposure with Europeans and thus these results support the idea that these sound-meaning relationships are not culture specific (Davis 1961).

The phenomenon Köhler discovered is known today as the bouba-kiki effect because of work done by Ramachandran and Hubbard (2001). Ramachandran and Hubbard (2001) showed English speaking participants and Tamil speaking participants a

jagged image and a rounded image and asked "Which is Bouba? Which is Kiki?" Over

95% of participants agreed that the angular image belonged with the word *kiki* and the

rounded image with *bouba*. The fact that both English speakers and Tamil speakers

demonstrated such a high level of agreement adds more evidence that this effect is not

restricted to one language group. Ramachandran and Hubbard argue that synesthetic

inter-sensory cross-connections are responsible for driving the bouba-kiki effect.

Synesthesia is the phenomenon where stimulation in one sense modality has an

automatic sensory experience in another sense modality. Ramachandran and Hubbard

(2001) argued that synesthesia-like effects might account for the bouba-kiki effect due

to motor cortical activation stemming from the mechanical act of articulation,

combining with the visual stimulation caused by the images presented. For example,

articulating *kiki* involves sharp inflections of the tongue and relates to the sharpness of

the jagged image and the rounding of the lips and oral cavity during the articulation of

bouba relates the roundedness of the images.

Ramachandran and Hubbard (2001) rely on an inter-sensory motor map of

articulation for their explanation of the bouba-kiki phenomenon; however, it is not

apparent that articulation is as important to listeners as the actual acoustic impact that

the sounds have (Ohala 1996). John Ohala (1996) argues that sound perception does not

necessarily involve retrieval of the motor pattern of articulation and emphasizes the

importance of how words sound as opposed to how they are articulated. Further, Ohala

(1996) emphasizes that infants and nonhuman animals, who are incapable of

articulating speech sounds, can nevertheless easily differentiate speech sounds. This

outcome suggests that motor articulation is perhaps not as important as the acoustic characteristics of the sound itself.

Maurer and colleagues (2006) investigated the bouba-kiki phenomenon with preschool children (age 2.6-2.10) to see if the effect would still occur in a relatively language inexperienced population. The preschool children were asked to participate in a game where they were read a story using stuffed animals as characters; they were subsequently asked to retrieve or pick an image or object for the stuffed animal character. After the story, the experimenter used the stuffed animal to ask the preschool children to respond to four typical bouba-kiki forced choice test trials involving matching a word with one of two images, one jagged the other rounded. Adults were asked to complete the study in much the same way as the preschool children, though the puppet was not used and they were informed that the task was designed for preschool children.

Maurer and colleagues (2006) statistically analyzed the choices made on the test trials in terms of whether the word had rounded vowels or non-rounded vowels; rounded vowels tend to be back vowels, such as /o/, and involve the rounding of the lips during articulation whereas unrounded vowels tend to be front vowels, such as /i/. They found that the rounded vowel words were matched with the rounded stimuli significantly more than chance and the non-rounded vowel words were matched with the jagged stimuli significantly more than chance. Interestingly, one trial involving the words goga and tiket were not statistically significant, suggesting that there could be

more than just differences in vowels driving the effect. This study is particularly

interesting because the preschool participants exhibited an effect, despite their relative

inexperience with language. These results are suggestive of an innate sound-sense

relationship.

Chris Westbury (2005) investigated the bouba-kiki phenomenon using an

interference task to find out if this effect is occurring on an implicit level and to test the

possibility that consonants are involved with the effect. Participants were presented

with a lexical decision task that involved determining whether they had seen a word or a

non-word. The words were presented visually inside of a frame that was either jagged

or rounded and the words themselves consisted of stop consonants (e.g. k, d, p),

sonorant consonants (e.g. m, l, n) or both stop consonants and sonorant consonants

mixed. The important dependent measure for such an interference task was response

time. Participants took longer to respond to the lexical decision task when the stop

consonant words were presented within the rounded frame and also took longer to

respond to the sonorant words when they were presented within the jagged frame.

These results support what Köhler (1947) and Ramachandran (2001) had demonstrated

previously and indicated that interference was occurring at a low-level.

Westbury (2005) conducted a follow-up study modifying the previously used

methodology so that, instead of using word stimuli, single letter stimuli were used. For

this follow-up study, participants were asked to indicate whether or not the character

displayed on screen was a letter. The results support the hypothesis that jagged frames

interfere with sonorant consonants and rounded frames interfere with stop consonants. In general, Westbury's (2005) study supports the notion that there is an implicit sound-sense connection between stop consonants and jagged images and between sonorant consonants and rounded images. Westbury (2005) suggests that these effects may be due to neurological cross-wiring in one of three candidate locations: the inferior temporal lobe and the angular gyrus, the left fusiform gyrus, and the left lateral posterior lobe.

The problem with Ramachandran (2001) and Westbury's (2005) synesthetic cross-connection accounts is that they fail to explain why the sound symbolic relationship should occur at all and in any specific direction; in other words, why are stop consonants associated with sharp imagery and sonorant consonants associated with rounded imagery and not the other way around? Although Ramachandran and Hubbard (2001) explain the bouba-kiki effect in terms of motor articulation and visual perception, it is not obvious what specific aspects of articulation lead to this association. One could argue that the quick bilabial opening and closing and flattening of the tongue during *bouba* could be related to the jagged image and the raising and rounding of the tongue during the articulation of the high-front vowels in *kiki* could be related to the rounded figure.

A central problem for sound symbolism more generally is: why do certain sound-sense meanings occur at all? Why does size symbolism associate high-front vowels with smallness and low-back vowels with largeness? Why are stop consonants

11

associated with jagged imagery and sonorant consonants with rounded imagery? What is the theoretical basis for sound symbolism?

1.3   Animal Communication: Implications for Sound Symbolism

Animal communication research offers a convincing theoretical basis for sound symbolic relationships. For example, Ohala (1994) posits that the reason high-front vowels are associated with smaller objects and low-back vowels are associated with larger objects is that, in general, larger animals produce sounds with a low fundamental frequency (F_0) and smaller animals produce sounds with a high F_0, which is an obligatory consequence of differences in the size of their vocal apparatus (see below). Ohala (1994) calls this "the frequency code" and suggests it can explain a number of sound symbolic relationships involving size that are also applicable to humans. For example, speakers who are unsure, polite or lack confidence will use a higher F_0 or a rising F_0; in contrast, speakers who are confident, assertive and authoritative will use a low F_0 or falling F_0 (Bolinger 1964, Bolinger 1978).

Ohala (1994) argues that "the frequency code" is biologically grounded, though it requires some experience and learning. Evidence of sexual dimorphism in voice characteristics supports Ohala's position (Fitch & Giedd 1999, Fitch & Reby 2001). For example, in humans the vocal folds of the larynx in males are approximately fifty percent longer than they are in females (Negus 1949). Further, the larynx is also descended slightly in the vocal tract of males compared to females, thus causing a 15-

20% increase in vocal tract length in males. Longer vocal folds and a longer vocal tract in males yield voices with lower F_0 and lower resonances, or formants. This sexual dimorphism arises at puberty and hence suggests that the voice differences between males and females are sexually selected perhaps reflecting attempts by males to sound larger or more threatening in the context of mate competition and mate choice (Ohala 1994, Fitch & Giedd 1999, Fitch & Reby 2001).

1.3.1 Motivation-Structural Rules

Morton (1994) points to other non-arbitrary relationships between sound structure and function. For example, he suggests that the structure of signals reflects the signaler's emotional state. For example, from analysis of signaling patterns in a variety of species, he observes that, in agonistic contexts, signalers generally emit a low-pitched harsh vocalization. In contrast, in contexts of fear or submission, signalers produce high-pitched, tonal vocalizations. Morton refers to these patterns as "motivation-structural rules" (M-S). The M-S perspective highlights two dimensions of signal structure: one is the quality of the signal that ranges from broadband and harsh, to narrowband and tonal; the other is the absolute frequency of the sound that ranges from low to high. A familiar example of M-S rules is the contrast between the growling and barking of an aggressive dog and the yelp or whimper of a submissive one. The "motivation" aspect of the model is that the structural variation in sound signals reflects the internal emotional states and motivation of the signaler.

There are some problems with Morton's model, however. For example, it is widely argued that signaling one's own internal emotional or motivational state is unlikely to be an evolutionary stable strategy in most contexts because it involves 'giving away' information about oneself that is potentially valuable to listeners and can be used against oneself (Dawkins & Krebs 1978). In most contexts, where the social and fitness interests of signalers and receivers conflict, even to a limited extent, selection should favour concealing rather than revealing such information.

This point highlights a broader problem in conventional signaling theory which has tended to model communication as being about transmitting information between a signaler and a receiver and where selection is presumed to favor reducing signal ambiguity so as to make communication more efficient (Dawkins & Krebs 1978). However, communication need not be about exchanging information at all. Assuming a broader evolutionary perspective, communication can be viewed as simply another means by which an organism can influence others (Dawkins & Krebs 1978). From this perspective, the signaler can be viewed as a self-interested actor that uses signals to manipulate and influence others to its own advantage. A dramatic example of how a signaler can manipulate the behaviour of other organisms is the common cuckoo, *Cuculus Canoris* (Davies, Kilner & Noble 1998, Dawkins & Krebs 1979). The cuckoo is well known for brood parasitism: the females lay their eggs in the nests of other species. The parasitized surrogate parents feed the cuckoo chick, despite discrepancies in size. Davies and colleagues (1998) found that cuckoo chicks manipulate parasitized reed warblers into feeding the cuckoo at an increased rate by producing rapid vocalizations that mimic

the calls of a brood of reed warblers. For example, when a blackbird or thrush chick was

placed in a reed warbler's nest accompanied by a loudspeaker, the provisioning rate of

the surrogate parents increased dramatically when a cuckoo begging song was played.

This rate of provisioning was equivalently matched when the vocalizations of a brood of

four reed warbler chicks begging was played. Although this example demonstrates vocal

manipulation between species, it could be argued that the begging calls of bird chicks in

a natural context use their vocalizations to manipulate their own parents into increasing

the rate of provisioning (Ottoson, Bäckman & Smith 1997).

1.3.2 The Affect Induction Model of Communication

A recent alternative model of animal communication that avoids the limitations

of Morton's M-S rules is the Affect Induction model of Owren & Rendall (1997). The

Affect Induction model follows the selfish-gene logic of evolution emphasized by

Dawkins and Krebs (1978) and argues that signaling is, first and foremost, a means of

influencing others in ways that benefit signalers and might, but need not, benefit

receivers as well. Owren & Rendall further emphasize that such influence is often likely

to involve exploiting low-level auditory and nervous system processes and processes of

arousal and motivation that are difficult for receivers to control or resist. They argue

that the acoustic signals are particularly suited to such modes of influence because they

are especially difficult for receivers to ignore or block-out. For example, simply turning

away or moving away can effectively reduce the effect of a visual or tactile signal, but,

for most species, one's ears cannot simply be shut-down or stopped-up in an analogous way to minimize the effects of acoustic signals.

One dramatic example of affect induction through the use of sound is that of "death screams": harsh sounding vocalizations with abrupt onset and high amplitude that prey animals, such as rabbits, exhibit during an attack (Wise, Connover & Knowlton 1999). Death screams are proposed to have an impact on the auditory and nervous systems of attackers in such a way as to induce an acoustic startle response that potentially enables the prey animal to escape (reviewed in Rendall et al. 2009).

The Affect Induction model has particular implications for animals that display dominance hierarchies. A low ranking animal that cannot defend itself in an agonistic context is predicted to produce vocalizations that have negative or aversive unconditioned effects on listeners, such as abrupt-onset and noisy, broadband shrieks and screams. Such aversive signals are proposed to be productive for signalers in such contexts by making continued harassment unappealing for the aggressor. In contrast, relatively dominant individuals are predicted to produce calls with prominent cues to individual identity and combine these with behavioral acts (aggression) with salient affective consequences for subordinates. The resulting conditioning effects should allow dominants to influence subordinates in the future through use of the calls alone without having to resort to costly aggressive acts (Owren & Rendall 1997).

The Affect Induction model of animal communication has predictions for the types of vocalizations used not only in agonistic or hostile situations but also in affiliative

or social situations. Once again, animals that display dominance hierarchies can benefit

by using vocalizations that elicit positive affect in conspecifics in those contexts where

close contact is desirable, such as during primate grooming. Such vocalizations can be

paired with other affiliative behaviours that strengthen the conditioned response and

thus strengthen the affective impact of the vocalizations. In general, this model predicts

that vocal sound patterns will change depending on what type of affect the signaler is

attempting to induce in a receiver. In most contexts, harsh and grating sounds are used

in conflict situations where they have perceptually aversive effects on listeners; while

smoother, more tonal and harmonic sounds are used in affiliative contexts where the

relatively harmonious qualities have an appeasing or soothing effect on listeners.

The Affect Induction model emphasizes the acoustic impact of vocalizations on

listeners but the affective effects of signals are not limited to vocal sound production

(Owren & Rendall 2001). A well-known example is that of the rattlesnake, aptly named

for its characteristic rattle located at the end of its tail (Fenton & Licht 1990).

Researchers have found that of the six species of rattlesnake studied, they all

maintained a characteristic broadband harsh sound pattern with rapid onset (Fenton &

Licht 1990). The authors conclude that the rattle is not meant for communication

between rattlesnakes because the signals are most intense outside the sensitive hearing

range of snakes; instead, the authors argue that this acoustic pattern has been selected

to instill fear in potential predators by startling them (Fenton & Licht 1990). This

example emphasizes the impact sound production has on listeners, regardless of how

the sound is made.

Although the Affect Induction model of communication has its origins in animal research, it has implications for human communication as well. The idea that sound can have an affective impact on human listeners is intuitive. A prominent example is that of music which is well-known for its influence on our emotional states; from heavy metal music that uses abrasive sounding vocals and instruments that can instill a sense of anger, to melodic classical music that can induce a sense of calm. Other every day examples include the negative effects of abrasive crying of human infants or the contagious positive effects of laughter (Rendall & Owren 2009). Such affective effects of sound are routinely exploited by the marketing and entertainment industries. For example, loud attention getting sounds are often used during television commercials, and many comedic programs use supplemental 'laughing tracks'. Such effects are exploited in even more draconian ways. For example, some militaries have researched and sometimes deployed acoustic signals (e.g., shrill, high pitched, high amplitude sounds) as weapons or crowd-control devices (Wilson 2010).

The Affect Induction model of communication could have other implications for the study of speech sounds in human languages. One particular area of interest is that of the sonority hierarchy, a ranking system of consonants and vowels from most vowel-like to the least vowel-like (Clements 1990). There are many different sonority hierarchies that have been developed over the years and have described the hierarchy in terms of relative intensity or amplitude, the openness of the vocal tract, and manner of articulation. Almost every hierarchy ranks vowels as the most sonorous followed by glides (/W/, /Y/), then liquids (/L/, /R/), then nasals (/M/, /N/) and finally obstruents

(/K/, /T/, /P/, /B/, /P/, /D/) are the least sonorous of all. Despite much work in this area of research, there is still no agreed upon theoretical basis for why sonority exists and should have the psycholinguistic impact that it does (Clements 1990, Parker 2002). This is precisely where an affect induction model of communication can play a role.

From an affect induction perspective, sonority could represent a progression from broadband harsh sounding consonants - analogous to animal vocalizations in hostile situations - towards harmonic voiced consonants and vowels, analogous to animal vocalizations in affiliative situations (Owren & Rendall 1997). A real world example of how these psychoacoustic properties of phonemes are pragmatically utilized is the use of the obstruent "shhh!" in an effort to quiet another individual or the use of the sonorant nasal "mmm" in the affiliative situation of eating food in a social context. Indeed, the most basic phonetic distinction between vowels and consonants may have its basis in this Affect Induction model of animal communication.

This thesis applied the Affect Induction model of animal communication to sound symbolism and human language. First, studies were conducted on the bouba-kiki phenomenon with emphasis on the consonant content of the nonsense words used. Specifically, words were generated using either plosive consonants (/T/, /K/, /P/) or sonorant consonants (/M/, /N/, /L/) in order to test the hypothesis that the psychoacoustic impact of the consonants used drives the corresponding jagged-rounded visual associations. The plosive consonants are characterized as having an unvoiced plosive burst of air pressure causing noise due to turbulence at the point of articulation,

while the sonorants are characterized as having a harmonic structure due to the vocal fold vibrations throughout (Ladefoged 2001). Thus, the plosive consonants were expected to coincide with the types of vocalizations used in agonistic contexts while the sonorant consonants were expected to coincide with the types of vocalizations used in affiliative situations.

Previous work on the bouba-kiki phenomenon has relied on hand-drawn visual stimuli and words that the researchers generated. To address the confound of possible experimenter bias in selecting stimuli, both words and images were generated using a random word generator for Study 1a and 1b, and a random image generator for Study 1a, 1b, and 2. Further, the methodology differed from previous work in that it did not entail presenting a rounded and a jagged image simultaneously; instead, participants chose between one of two words for a given single image in Study 1a and 1b and in Study 2 participants actively morphed an image by adjusting a sliding scale from rounded to jagged for a single word. Arguably, these methodological changes allow for a more systematic and less biased investigation of the bouba-kiki phenomenon. Finally, an analysis of English swearwords, heavy metal lyrics, carols and lullabies was conducted to buttress the Affect Induction model's applicability to the real world. The swearword/profanity list and heavy metal lyrics were hypothesized to contain phonemes that rank low in sonority in order to gain the most from negative affect induction. On the other hand, lullabies and carols were predicted to contain phonemes that rank high in sonority in order to induce positive affect or to convey joy.

CHAPTER TWO

2.1 Introduction

Sound symbolism is a difficult subject to study because of the confounds related to actual word usage and language expertise. To address the issue of word usage and language expertise, experimental designs using unfamiliar stimuli are required. The bouba-kiki phenomenon may provide an avenue for experimental designs that use unfamiliar stimuli. The bouba-kiki phenomenon has typically been investigated by presenting two nonsense words and two images simultaneously to a participant and asking him or her to match the words with the images. The images are typically generated by the experimenter by hand and vary in that one image is angular and jagged while the other image is rounded and smooth. Further, the words are also generated by the experimenter. For example, Köhler (1949) originally used the words *takete* and *baluba* and later used *maluma* instead of *baluba* and Ramachandran and Hubbard (2001) used the words *kiki* and *bouba*.

There are several reasons why this methodology is not as robust as it could be. First, experimenters generating their own word and image stimuli risk introducing biases into the stimulus materials by effectively selecting words and images that they unwittingly feel go together. When participants subsequently match images and words correctly, they may then simply be confirming the original biases. Second, when both types of images (rounded and jagged) and both types of words (plosive and sonorant) are presented simultaneously, the task may become too transparent to participants

21

because they can explicitly compared the differences in images with the differences in words and develop a sorting rule that they implement on all trials.

The objective of Study 1 was thus to modify the conventional bouba-kiki methodology to try to reduce these biases and transparency issues. Potential biases in experiment's selection of stimuli were reduced by utilizing randomization techniques to create both image and word materials. Participant biases were reduced by presenting individuals images serially rather than simultaneously in pairs.

2.2 Experimental Designs Overview

Study 1a consisted of a two alternative forced choice (2-AFC) procedure involving matching one of two words (plosive or sonorant) with a single image; study 1b replicated study 1a with a new set of images. Study 2 differed in allowing participants to morph the shape of individual images to a form they thought best suited a single word presented to them.

2. 3 Study 1: Forced Choice Method

2.3.1 Participants

Participants in Study 1a were 41 undergraduate students (20 Female, 21 Male) enrolled in introductory psychology courses at the University of Lethbridge and they

received partial course credit for their participation. Participants in Study 1b were from the same subject pool as Study 1a, except that all participants were female (n = 21).

2.3.2 Stimuli

Image Generation:

An ideal solution to avoid possible experimenter biases in image generation would be to generate stimuli by way of a random image generator that can be mathematically manipulated to progressively change from an angular image to a rounded image. One such program is the "random shape generator" (Neil Birkbeck 2008) which uses Bézier curves to alter an image from a piecewise linear outline to a progressively rounded image. Bézier curves are mathematical functions that alter the path of a line by the simultaneous influence of control points. Bézier curves have been used in automotive design to manipulate the design of car body lines and have recently been used in scalable vector graphics.

In the current study, image seeds were randomly generated using this "random shape generator" program (Neil Birkbeck 2008). The random shape generator generates a set of random initial calculus points to form the base image or "seed". Each of these initial points, also known as interpolated points, have two control points that alter the arrangement of the shape through the use of cubic Bézier curves. Using a radially constrained methodology, randomly generated points populated a finite space and

were joined via a post-process to form an array of points that make up the seed image. Each seed image generated contained eight separate initial points that, when directly connected, form the 0% smooth image (i.e. a jagged image). Further, each image seed can generate images ranging from 0% to 100% "smoothness". With each incremental increase in "smoothness", the points migrate from an initially jagged configuration to a progressively rounded configuration; this was done by progressively moving the control points away from the interpolated points. An intuitive way of thinking about Bézier curves is that each interpolated point is "pulled" by the control points and the farther the control points are away from the interpolated points the stronger their pull. It is important to note that because there is no agreed upon method of objectively measuring "smoothness", "percent smooth" is not an actual percentage of smoothness in any absolute sense; instead, "smoothness" should be interpreted in relative terms as the proportion of incremental change that the control points exert on the Bézier curve.

All images were black line figures with a white background, some of which were self-intersecting. In the current study, 20 seed images were generated, however, only 0%, 15%, 30%, 50%, 70% 85% and 100% smooth images were used, resulting in 140 images in total (seven per seed). Images were saved as bitmap files at a 480x480 resolution. To see an example of a series images ranging from 0%-100% smooth refer to Figure 2.1. Image generation for Study 1b was identical to Study 1a, except a new set of 20 seed images was generated and used.

Word Generation:

Words were generated using a random word generator program called the Gammadyne Random Word Generator. This program generated two syllable words consisting of four letters in a consonant-vowel, consonant-vowel (cVcV) format (e.g., Mula or Tiki). Experimental words contained either entirely plosive consonants (/T/, /K/, /P/) or entirely sonorant consonants (/N/, /L/, /M/). The plosive consonants are perceptually harsh sounding and are unvoiced, meaning that they occur without vocal fold vibration. The sonorant consonants are perceptually smooth sounding and are accompanied by voicing which gives them a more harmonic structure. Locus of articulation was partially controlled: thus /T/ and /N/ are both alveolar-dental consonants, /L/ and /K/ are both approximately palatal consonants and /P/ and /M/ are both labial consonants. The vowels used were /E/, /A/ and /U/. /E/ was restricted to occur only as the first vowel because an /E/ as the ending of a word could shorten it to one syllable (eg. Pate, Nule). Vowels /I/ and /O/ were not included for word generation due to possible confounds from size-sound symbolic effects associated with these vowels (Sapir 1929). In total, 40 words were generated, 20 plosive and 20 sonorant.

This corpus of words was further grouped into four different categories of five word pairs each based on various additional controls they represented. Thus, five pairs of words controlled for vowel content by using the same vowels in both the plosive and sonorant word pair (e.g. Ketu, Melu). Five word pairs controlled for locus of articulation matched (e.g. Paku, Mala), although, for two word pairs, this could only be partially

controlled (Keta-Lemu; Taku-Nemu). Five word pairs controlled for both vowel content and locus of articulation. Finally, five word pairs were generated entirely randomly with respect to vowel content and locus of articulation. For a complete list of words used in study 1a and 1b refer to Table 2.1.

2.3.3 Procedure:

The study was conducted on a Dell desktop computer running Windows XP in a sound isolated room. Participants were asked to read instructions about the study and informed consent was obtained. As part of the written instructions, participants were told that the words they were about to read came from a dialect of Calacotan native to Polynesia. This was done in order to alleviate potential confusion over the unusual nature of the words used. Participants responded via a graphical interface created by Runtime Revolution v. 2. 9. In each trial, the participant was shown a line image and below this line image the participant was shown two side-by-side buttons with a word on each of them, one plosive and one sonorant. Every participant received the same plosive and sonorant word pair for a given image but the right-left placement of the words and the order of trials was randomized across participants.

Participants were then asked, via written instructions, to complete the task as follows: "your job is simply to select which of the two words you think is more appropriate for the drawing and press that button using the mouse". Selecting a word button automatically advanced the participant to the next trial. Between trials a black

and white fixation cross appeared on screen for one second to signal the next trial.

Participants were given eight seconds to complete each trial. After eight seconds,

participants received a "timeout" message indicating that they had taken too long to

respond. Timeouts ultimately accounted for less than 1% of trials for both study 1a and

1b. To see an image of what a trial looked like to participants, refer to Figure 2.2.

2.3.4 Results:

A mixed between and within-subjects analysis of variance (ANOVA) was

conducted on the dependent measure of sonorant versus plosive word choice as a

function of the between-subjects factor of sex and the within-subjects factors of image

smoothness and word control group (vowel matched; locus matched; locus and vowel

matched; and random). There was no main effect for sex or for word control group ($F(1,$

$39) = 2.29$; ns, $F(3, 39) = 1.18$; ns). There was, however, a main effect for image

smoothness ($F(6, 39) = 43.12$; $p < 0.01$). There was no interaction between sex and

image smoothness ($F(6, 39) = 0.11$; ns). A Fisher's least significant difference (LSD)

multiple-comparison analysis was also conducted on whether a plosive or sonorant

word was chosen as a function of image smoothness (0%, 15%, 30%, 50%, 70%, 85%,

100%). 0%, 15% and 30% smooth images led to less sonorant word choices and were

different from all other images, however, 50%, 70%, 85% and 100% smooth images did

not differ from one another (Figure 2.3).

To assess the vowel content of words another ANOVA was conducted for word choice as a function of sex, image smoothness and vowel content. A filter was applied to only include word pair groups that controlled for vowels: vowel matched and vowel and locus of articulation matched. There was no main effect for sex or for vowel content ($F(1, 39) = 1.92$; *ns.* $F(5, 39) = 1.36$; *ns*); however, there was still a main effect for image smoothness ($F(6, 39) = 26.15$; $p < 0.01$).

In order to replicate Study 1a, similar analyses were conducted on the dataset from Study 1b. The only exception is that the sex factor was not included in these analyses because Study 1b involved only female participants. An ANOVA was conducted for word choice as a function of image smoothness and word control group (vowel matched, locus of articulation matched, vowel and locus matched and random). There was no main effect for word control group $F(3, 20) = 1.63$; *ns*); however, there was a main effect for image smoothness $F(6, 20) = 20.69$; $p < 0.05$. A Fisher's LSD post-hoc multiple-comparison analysis was conducted on image smoothness. 0% smooth images differed from every other image smoothness. 15% and 30% smooth images differed from all other images but they did not differ from one another. 50% and 70% smooth images differed from 0%, 15%, 30% and 100% smooth images but did not differ from one another or from 85% smooth images. 85% smooth images differed from only 0%, 15% and 30% smooth images. 100% smooth images differ from all other images except for 85% smooth images (Figure 2.3). In order to investigate the possible influence of vowel content, another ANOVA was conducted for word choice as a function of image smoothness and vowel content. There was no main effect for vowel content $F(5, 20) =$

1.51; *ns.*). There was, however, a main effect for image smoothness $F(6, 20) = 22.24$; $p <$ 0.01).

2.3.5 Discussion:

The prediction that the nature of consonants (plosive vs. sonorant) has an impact, or perhaps even drives the bouba-kiki phenomenon, is supported by the current study. If words are arbitrary and their acoustic patterning inconsequential, word choice selection should be at chance rates with respect to image smoothness. The main effect between word choice (plosive vs. sonorant) and percent smoothness provides evidence for sound symbolism. The lack of a main effect for vowel content suggests that consonant content had a greater effect on word matching decisions. Figure 2.3 depicts a positive relationship between word choice and image smoothness, such that as the image becomes increasingly rounded, it is more commonly matched with sonorant words. Figure 2.3, for Study 1a, also demonstrates a plateau at 50%, 70%, 85%, 100% smoothness and these points are not different from one another. This possibly indicates that the image may become sufficiently rounded at 50% smooth that additional roundedness is not perceptually salient.

Study 1b replicated the findings of Study 1a and because new images were used, this indicates that the results from Study 1a were not due to idiosyncratic associations specific to those image and word pair combinations. Study 1b had only female participants and this potentially limits its external validity; however, Study 1a involved

both male and female participants and found no difference in their word-matching

performances. Study 1b also replicated the non-significance of vowel content when

using a filter to include only the word-pair image combinations that controlled for vowel

content. Although these findings suggest that vowel content is less important than

consonant content, they do not rule out vowel effects completely nor do they say

anything about vowels that were not tested. Additional research is necessary to further

test what role vowels play in such word matching tasks. The multiple comparison post-

hoc tests indicate that the relationship between average word choices across image

smoothness differed between Study 1a and 1b. Study 1a appears to plateau after 50%

and Study 1b appears to have a somewhat more linear relationship.

2.4 Study 2: Slider Method

The objective of this study was to try to further deconstruct the bouba-kiki

phenomenon and to address some of the methodological weaknesses of Study 1. This

was accomplished by allowing a single image to be morphed by the participant to a

shape they felt best matched an experimental word. This modification introduces some

flexibility into subject responses and overcomes the limitation that previous effects may

have been driven by a limited choice of exaggerated shape options. An additional

potential weakness of Study 1 was the possibility of orthographic effects, specifically

that visual jaggedness or roundedness of the letters used to make the words could

influence how participants matched them to the jaggedness or roundedness of the

images presented. Study 1 attempted to control for this orthographic effect by using sonorant consonants that are jagged in uppercase form (M, N, L) and a plosive that has a distinct roundedness (P). Despite this, visual associations between the word stimuli and the images used could still be a factor, especially for uppercase and lower case /T/s and /K/s. In order to control for this possibility, Study 2 presented the word stimuli aurally through a set of headphones using a voice synthesizer to produce the words.

The predictions for Study 2 mirrored the predictions for Study 1: words containing plosive consonants were predicted to be matched with jagged images, while words containing sonorant consonants were predicted to be matched with rounded images.

2.4.1 Participants:

Participants were 32 undergraduate students (16 Female, 16 Male) enrolled in introductory psychology courses at the University of Lethbridge and they received partial course credit for their participation. Two additional participants were dropped from the analysis because their reaction times were dramatically lower than the average ($M = 3.49$ seconds) for all other participants suggesting that they were not attending to and manipulating the image stimuli earnestly.

2.4.2 Stimuli:

Image Generation:

Images used in Study 2 were identical to those used in Study 1, except that every

incremental increase in image smoothness was available to subjects as they morphed

the images themselves. Because participants' performance in Study 1a suggested a

possible smoothness plateau effect at around 50% smoothness, the range of

smoothness increments available to participants in this experiment was truncated at

50%. In addition, only seed images that contained line-segments that formed acute

angles (< 90 degrees) were used. This constraint was included to make the images more

comparable and to address the possible confound of having seed images that are

already relatively rounded to begin with because they include a large number of obtuse

angles (i.e., in their more most jagged form they appear comparatively rounded). In

total, there were 144 image seeds used each with 51 possible smooth increments

available (0-50% smooth). Images were saved as bitmaps at a resolution of 480x480.

Word Generation:

Word stimuli were generated by taking a subset of English consonants and

vowels and forming all possible consonant-vowel consonant-vowel (cVcV) combinations.

The consonants and vowels used were identical to Study 1a and 1b except that the

vowel /E/ was dropped to reduce the number of possible combinations and to avoid

changing words into a single syllable (e.g. Pule, Kute). When the word generation

process generated what proved to be a real word (e.g. Papa or Tuna) these were

retained but flagged for later analyses. Mixed words, i.e., those containing both a

plosive and a sonorant consonant, were included in this study. There were 36

consonant configurations and four vowel configurations, yielding a total of 144 word

stimuli. Words were then entered into SwiftTalker, a commercial text-to speech voice

synthesis program by Cepstral. A variety of voices can be synthesized using SwiftTalker

that vary in the sex of speaker, accent and age. The voice used in this study was "David",

an adult male speaker of American English. All other options were left at their default

settings and the synthesized words were saved as WAV files.

2.4.3 Procedure:

The general materials and procedures used for this experiment replicated those

from Study 1a and 1b. Demographic information and informed consent was collected

from participants; additionally, participants were asked if English was their first

language. Following this, participants selected an "instructions" button on the screen

that informed them that they were about to participate in a study that involved

changing an image to match a word they would hear through headphones (Sennheiser

HD 280 pro). Participants were subsequently instructed to select the "Demo" button to

watch an audio-visual demonstration of a trial. During the demonstration trial,

participants received written instructions on screen and watched an image morph as

they heard a nonsense word produced three times. They were then instructed to move the slider in order to morph the image to best match the word that they heard. After the demonstration trial, participants selected a button to begin test trials.

All trials began with a black and white fixation cross presented for one second. After the fixation cross, participants were presented with an image stimulus that continuously morphed through the complete range of variation to illustrate the shape possibilities. At the same time, the participants were aurally presented with the word stimulus three times. In order to avoid biasing participants' subsequent morphing responses, the image began in the middle (25% smooth) then moved to one extreme (0% or 50% smooth) then back through the midpoint to the other extreme. The nonsense word for that trial was played three times over the headphones and each time at the point when the morphing image was at midpoint. This demonstration occurred at the beginning of every trial and the direction the morph moved was randomized so that approximately half of the time it initially went to the 0% smooth extreme and the other half of the time it initially went to the 50% smooth extreme. Throughout the audio-visual demonstration, a slide bar moved in concert, such that at 50% smooth the slide bar is at its furthest right position and at 0% smooth the slide bar is at its furthest left position. Upon completion of the audiovisual demonstration, participants were then able to morph the image by moving the slide bar to best match the word they heard. After making their selection, participants then selected a "next trial" button. Also included was a "replay word" button that participants could select to replay the word after the demonstration phase of each trial. Each trial had a 15 second time limit after

which the participant received a "timeout" and a message encouraging faster replies. Timeouts ultimately accounted for less than 1% all participant trials. See Figure 2.4 for an example of the visual layout of a trial.

2.4.4 Results:

A mixed effects analysis of variance (ANOVA) was conducted on the dependent measure of slide-scale choice (0-50% smooth) as a function of the between-subjects factor of sex and the within-subjects factor of word type (plosive, sonorant or mixed). There was no main effect for sex and no interaction between sex and word type ($F(1, 30) = 1.05$, *ns*, $F(2, 30) = 0.56$, *ns*.); however, there was a main effect for word type ($F(2, 30) = 17.19$; $p < 0.01$; Fig 2.5). Fisher's LSD post-hoc tests indicate that plosive words were significantly different from sonorant words and mixed words were significantly different from both sonorant and plosive words.

To test whether or not real words (e.g. Papa, Tuna) influenced the findings, the same ANOVA was conducted with a filter to include only nonsense words. The word type variable still had a main effect ($F(2, 30) = 15.24$; $p < 0.01$) and there was still no main effect of sex ($F(1, 30) = 0.87$, *ns*). For this reason no filter was used removing the effects of real words for the rest of the analysis. To investigate a possible vowel influence, an ANOVA was conducted using the dependent measure of slide-scale choice as a function of sex and vowel content. There was no main effect of vowel content or sex ($F(3, 31) = 0.94$; *ns,* $F(1, 30) = 1.43$; *ns*).

In order to analyze 'pure' words that used only one consonant type, an ANOVA was conducted using the dependent measure of slide-scale choice (0-50% smooth) as a function of sex and sonority word type with a filter to include only words with the same two consonants (i.e. K K, P P, T T, L L, N N, M M). There was no main effect for sex nor for a sex and word type interaction ($F(1, 30) = 0.12$, *ns*, $F(2, 701) = 0.00$, *ns*); however, there was a main effect for word type $F(1, 30) = 11.10$; $p < 0.01$), indicating that there was a difference between plosive and sonorant word types with sonorant words eliciting a greater smoothness selection by participants. To better compare specific consonants directly, another ANOVA was conducted using the dependent measure of slide-scale choice as a function of consonant content with a filter to include only words with the same two consonants. There was a main effect for the consonant content variable ($F(5, 30) = 3.54$; $p < 0.01$, Fig 2.6). Fisher's LSD post-hoc tests indicate that words with two Ks and words with two Ts were significantly different from words with two Ns, words with two Ms and words with two Ls. However, words with two Ps were not significantly different from other same two letter words. This suggests that P may not be as harsh sounding to participants as T and K.

2.4.5 Discussion:

If nonsense words are indeed arbitrary and void of meaning then it would be expected that there should be no significant difference among words. From these results, however, it is clear that this null hypothesis can be rejected. For example, the

main effect for word type indicates that there are significant differences between plosive, sonorant and mixed words. The overall mean for plosive words was below that for sonorant words and the mean for mixed words was in between; these overall results coincide with predictions based on the Affect Induction model. Further, the lack of a main effect for vowel content, combined with the significant main effect for consonant content, suggests that consonant content is a more important factor than vowels. Of course, this study only utilized two vowels (U, A) and so additional studies using a larger set of vowels are required to more fully evaluate the influence of vowels on object-shape, word choice matching tasks like these.

The mean slider response value was expected to differentiate the consonants into a hierarchy with the plosives clustered near the 0% smooth end and the sonorants clustered near the 100% smooth end. From Figure 2.5 it is clear that these differences were not as pronounced as hypothesized. Nevertheless, the ordinal rank order of the different consonants is precisely what would have been expected based on our hypothesis that harsher sounding plosive consonants will be more commonly associated with jagged imagery and smoother sounding sonorant consonants will be more commonly associated with rounded imagery.

Another limitation of both Study 1 and Study 2 is that they did not use all possible consonants and vowels. This limitation is difficult to avoid in part because a complete set of consonant and vowels would have entailed too many word combinations to sample completely thus yielding participant response fatigue. At the

same time, high-front vowels (such as /i/) and low-back vowels (such as /o/) have

additional size-symbolic connotations (Sapir 1929, Ohala 1994, Ultan 1978) that might

interact with and thus confound the hypothesis about image smoothness being tested

in the current experiments. Hence, additional studies are required to more fully

evaluate the influence of other consonants and vowel combinations.

2.4.6 An Affect Induction Explanation of the Bouba-Kiki Phenomenon

Although emphasis has been placed on the phonemic and acoustic aspects of

affect induction, this does not mean that affect induction only occurs in the auditory

domain. The Affect Induction perspective of signaling can be applied to all sense

modalities as all senses can be manipulated or influenced (Owren, Rendall & Ryan

2010). Many animals, for example, produce signals that affect the visual, olfactory,

gustatory and somatosensory systems of conspecifics and predators. For example, cats

in threatened situations will not only hiss but they will also arch their backs, bare their

teeth and appear larger due to the piloerection of their fur. Some animals, such as

skunks, produce noxious chemicals that affect the olfactory system and other exposed

sensitive tissue. Even physical strikes can be considered part of signaling, according to

an affect induction perspective. The use of these multimodal signals presumably affects

the senses in such a way as to cause affective changes.

Multimodal affect induction may have explanatory implications for the bouba-

kiki phenomenon. For example, shapes that vary in terms of smoothness have been

shown to have different effects on participants (Bar & Neta 2006, Bar & Neta 2007, Aronoff 2006). For example, participants in one study were briefly presented with an image that was either rounded or angular and were asked to rapidly respond with either "like" or "dislike" based on their "gut" feeling (Bar & Neta 2006). The images were either of real world objects such as a couch or they were "meaningless" patterned images and each image had both a rounded and jagged version. Participants generally preferred the rounded version over that of the jagged version for both real world objects and for the "meaningless" patterns (Bar & Neta 2006). The authors of this study argued that when making rapid judgments participants interpreted jagged objects as more threatening or dangerous than their rounded counter-parts and that this threat assessment is very low level and potentially nonconscious (Bar & Neta 2006). Bar and Neta (2007) followed up this study by analyzing the effects of jagged and rounded shapes on participants while having their cortical activity monitored through functional magnetic resonance imaging (fMRI). Participants had greater activation of the amygdala, an area of the brain implicated with fear response, for sharp or jagged imagery compared to rounded imagery (Bar & Neta 2007). These findings coincide with the prediction that images with sharp angular contours elicit a low level threat response.

Further evidence of the relationship between affect and shape is provided by Aronoff and colleagues (Aronoff 2006). For example, Aronoff and colleagues (1988 as cited in Aronoff 2006) asked participants to draw a mask for a hostile situation of war and a mask for an affiliative situation of courtship. After analyzing the types of shapes participants drew, the authors found that the masks drawn for hostile situations

consistently had jagged features such as pointed eyebrows, horns, teeth and angular

beards (Aronoff, Barclay & Stevenson 1988 as cited in Aronoff 2006). This contrasts with

the types of shapes participants drew for the mask meant for an affiliative situation

which were rounded and curvy. A follow-up cross-cultural analysis of masks also

supports the use of jagged features for hostile masks and rounded features for affiliative

masks (Aronoff et al. 1988 as cited in Aronoff 2006). Aronoff and colleagues (1992 as

cited in Aronoff 2006) have even found an effect for the angularity of movements of

ballet dancers and an association with threatening ballet characters as well as the

rounded movements of ballet dancers and an association with warm and affectionate

ballet characters.

It is my contention that the bouba-kiki effect is explained by the corresponding

affect induced in the visual and auditory domains: the negative affect induced by the

jagged imagery corresponds with the negative affect induced by the harshness of the

plosive consonants and the positive affect inducing rounded imagery corresponds with

the smoothness of the sonorant consonants. Both the visual and auditory aspects of the

stimuli involved in the bouba-kiki effect activate a low-level threat assessment via the

limbic system and allow a common valence of affect to form associations among

otherwise nonsense and random stimuli.
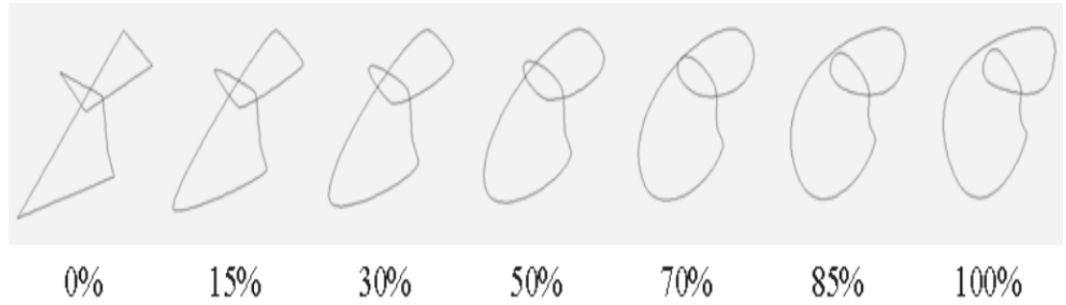
Figure 2.1. Example of a seed image changing from initial linear configuration to progressively rounded configurations at 0%, 15%, 30%, 50%, 70%, 85% and 100% relative smoothness. The image is darkened here to improve contrast; the background was white during the actual study.
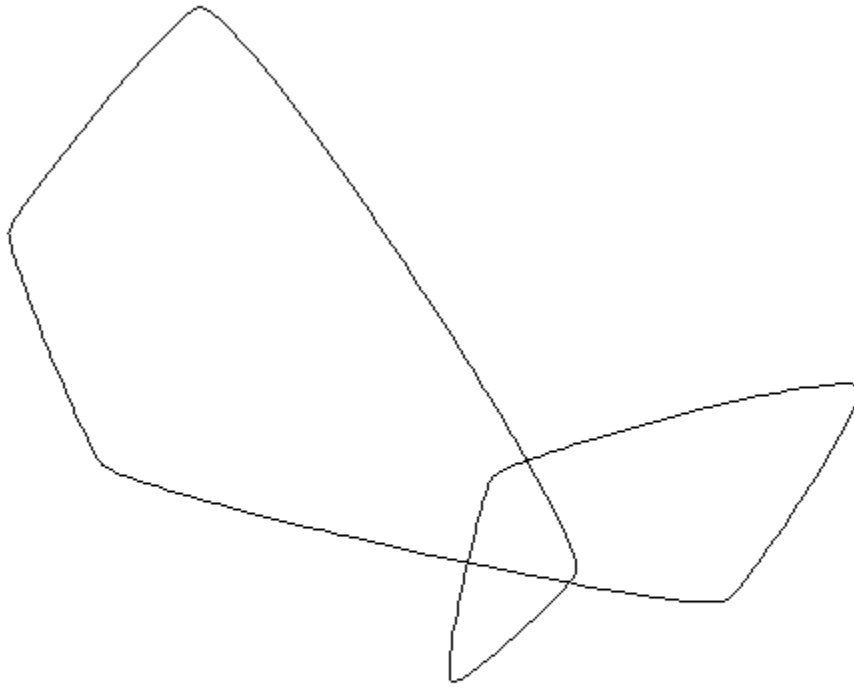
Figure 2.2 An example of a trial from Study 1a.

Word Choice Across Image Smoothness
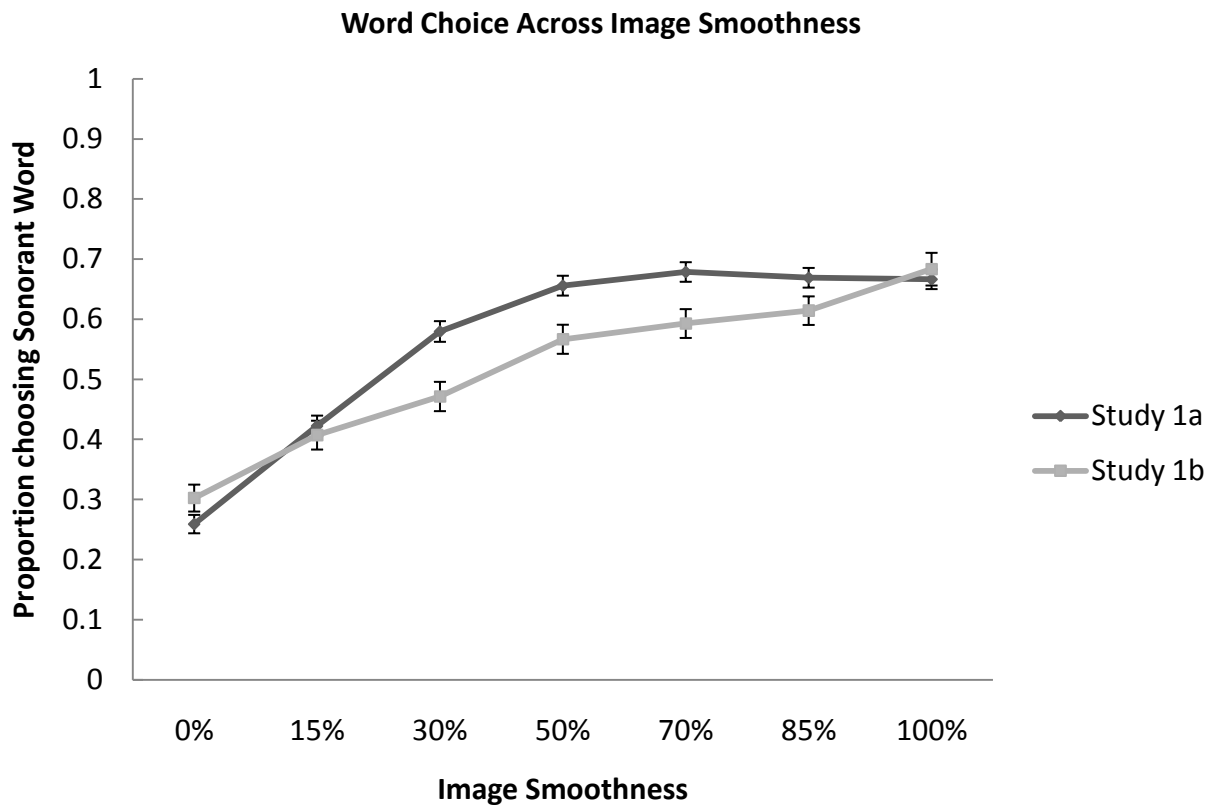
Figure 2.3. Mean response of choosing either the sonorant word or the plosive word across images of varying smoothness (0%, 15%, 30%, 50%, 70%, 85%, 100%) for studies 1a and 1b. Sonorant words were assigned as a dummy variable of 1 and plosive words were assigned a dummy variable of 0. Error bars indicate standard error.
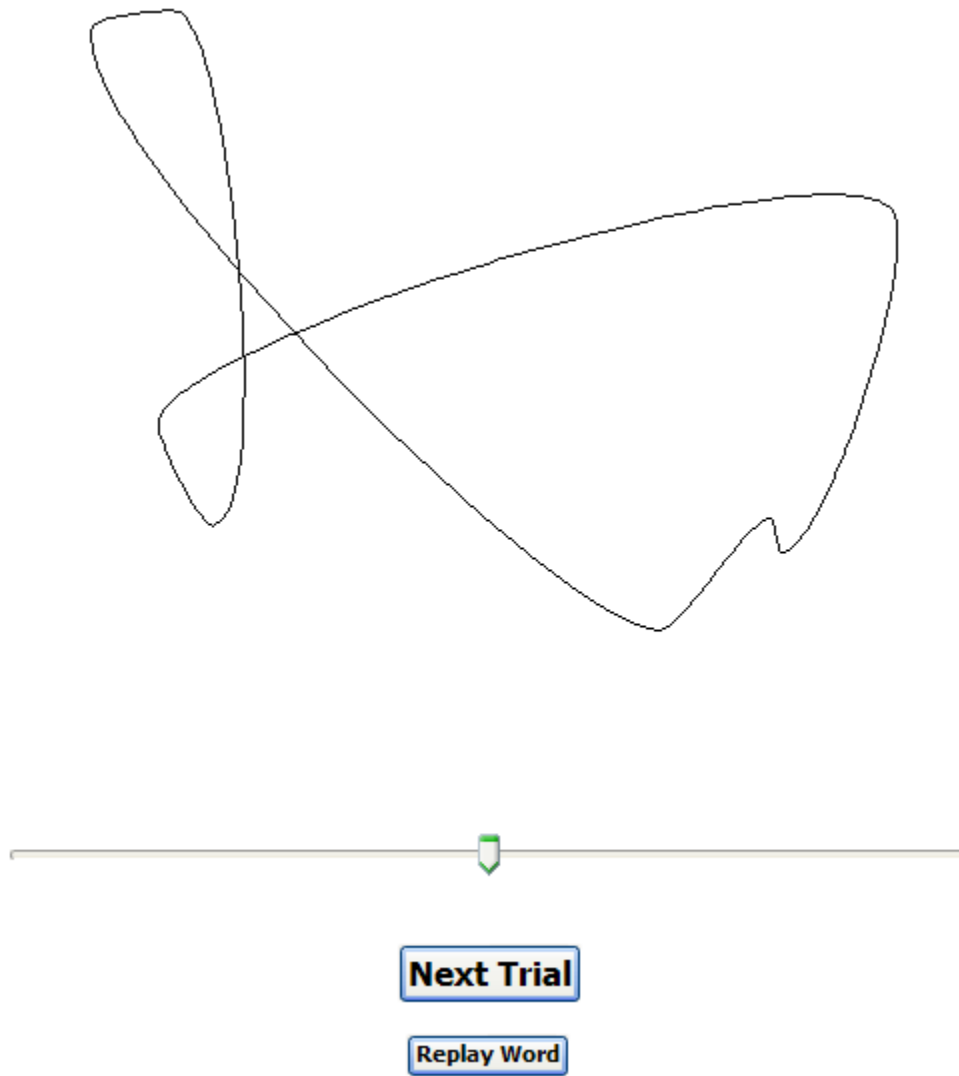
Figure 2.4. An example of a trial from Study 2.
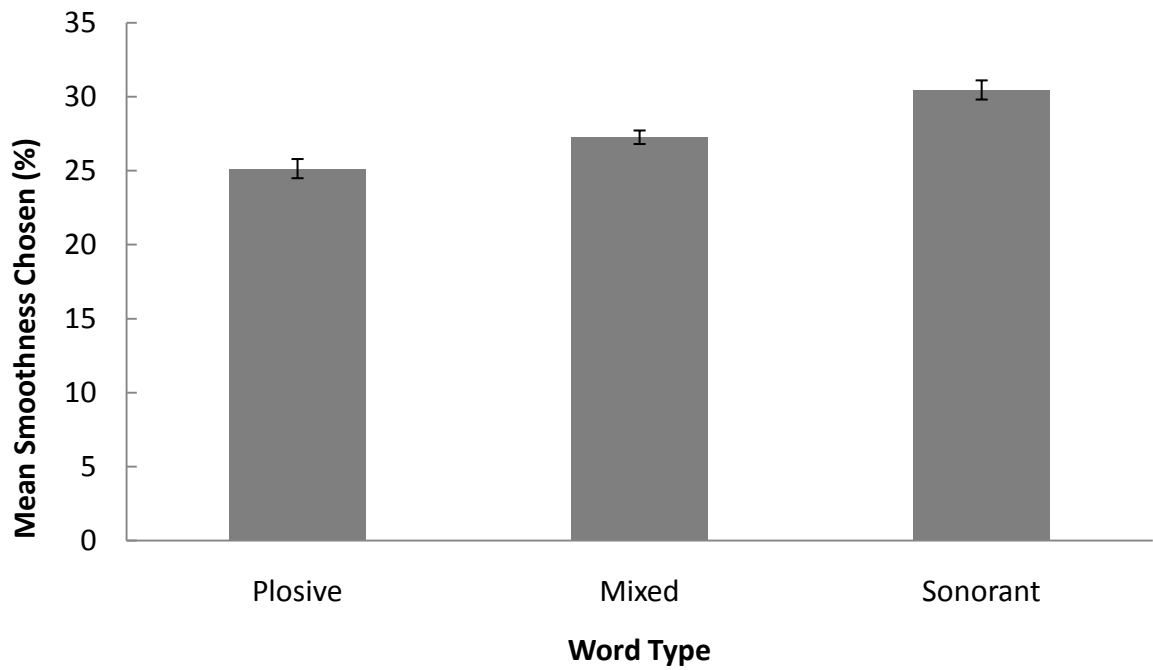
Figure 2.5. Mean smoothness chosen as a function of the sonority-type of word given. Error bars indicate standard error.
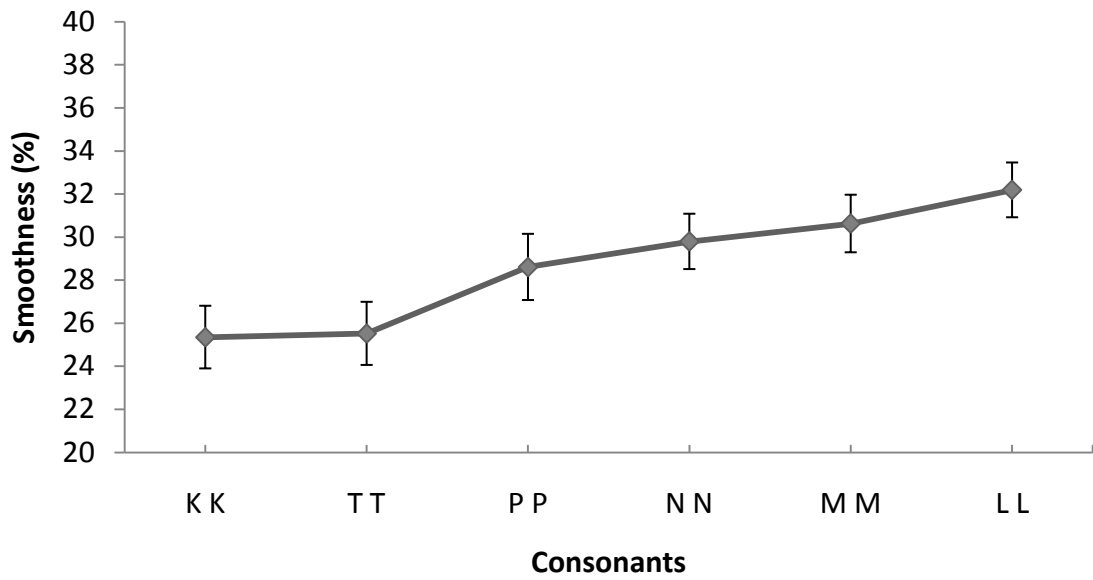
Figure 2.6. A line graph depicting the mean smoothness chosen for words with the same two consonants. Error bars indicate standard error.

Vowel matched

Plosive	Sonorant
Pata	Nala
Peku	Lenu
Ketu	Melu
Keta	Lema
Tapu	Manu

Consonant locus of articulation matched

Plosive	Sonorant
Keta	Lemu**
Kepu	Luma
Taku	Nemu**
Paku	Mala
Petu	Mena

Vowel and locus of articulation matched

Plosive	Sonorant
Teku	Nelu
Kutu	Lunu
Kapa	Lama
Puka	Mula
Tepa	Nema

Random

Plosive	Sonorant
Keka	Memu
Kapu	Nenu
Putu	Nela
Pepa	Muna
Tuka	Lamu

Table 2.1. Study 1a and Study 1b list of plosive and sonorant word pairs in four groups: vowel matched, locus of articulation matched, vowel and locus of articulation matched and random. ** Lemu should be Lenu and Nemu should be Nelu to fully approximate locus of articulation.

CHAPTER THREE

3.1 Introduction

The findings presented in chapter two suggest that there could be some relationships between the sounds of words and the kinds of objects they naturally signify mediated through the types of consonants involved in the words: nonsense words with harsher sounding plosive consonants were associated with jagged imagery while words with smoother sounding sonorant consonants were associated with the rounded imagery. These findings support the Affect Induction model; however, they do not tell us that such relationships are manifest in real words. To test this possibility, in this chapter I examine a corpus of real words for evidence that the sound-meaning relationships identified in Study 1 and 2 extend to real words. The analysis focuses on samples of real words used in categorically distinct genres of literature or music whose themes are hypothesized to reflect distinctly different affective tones and thus where any word-form relationships that signal affective tone would be expected to exert an influence. As a first pass at such a test, I selected as examples of positively toned genres, children's lullabies and carol-type songs. Lullabies are stories or songs narrated or sung by parents with the specific objective of soothing and pacifying infants and young children particularly at times when they are fussing or distressed. Hence, they are a clear example of the use of sound to influence the affective state of listeners. Likewise, carols are generally cheerful songs used to instill positive affect during the holiday season. They are intended to be uplifting, happy and celebratory.

As examples of negatively toned genres to contrast with lullabies and carols, I selected swearwords (profanity) and lyrics from heavy metal songs. Swearwords and profanities are often used in aggressive circumstances and are conventionally intended to cause offense and to induce negative affect. Of course, profanity can be used in a much wider range of circumstances, including for comedic effect or simply to shock a listener, and there are a variety of situational factors that will determine the quality of affect it might engender (Jay & Janschewitz 2008). Nevertheless, in broad terms, profanity provides an intuitive contrast for lullabies, where the affect-inducing effects of word-forms would be predicted to vary categorically: positive or calming (arousal reducing) in the case of lullabies; negative or exciting (arousal increasing) in the case of profanity. Heavy metal music contrasts with carols in being moody and conveying anger, hostility or other negative affect.

Given these dichotomous associations, I predicted that, if the sound-meaning relationships explored to this point were preserved to any degree in real words, that lullabies and carols would contain a preponderance of words with sonorant consonants, while profanity and heavy metal songs would contain a preponderance of words with harsh plosive consonants.

3.2 Methods

Word Selection

All words were obtained from online sources and were located by search engine results based on key words "lullabies", "carols", "swearwords list", and "heavy metal lyrics". Twenty-five lullabies were randomly selected from a website containing a corpus of 163 lullabies (Kluytmans 2010). Twenty carols were randomly selected from a corpus of 59 carols (Smith 2010) and twenty heavy metal songs were selected from a heavy metal lyric website containing over 12,000 albums (McDonnell 2010). To select the heavy metal songs, a random number was generated for the letter of the alphabet and then a random number was generated to select an artist, then album and finally song. This randomized selection process was done to avoid selection biases as far as possible. All random numbers were generated using an online random number generator (Haahr 2010). Finally, the profanity words were selected by combining word lists from a variety of profanity filters and lists found online. These word sources included: an internet video of an unofficial list of profane and offensive words that are optionally censored in the comments section on YouTube; a swearing/profanity web site; a profanity list used by a "politifier" program designed to replace swearwords with euphemisms; and a profanity list meant for webmasters and forum administrators (Maxwell 2010, Jones 2010, Back 2010, Thompson 2010).

To minimize possible bias in word selection I did as little filtering of material as possible. Many of the words in the swearing/profanity list were sexual in nature and

could arguably have been removed to limit the list to only pejorative swearwords; however, many insults and swears involve sexual themes and thus sexual profanity was not removed to avoid introducing some bias through this kind of subjective judgment. All variants of words were also retained across lists in order again to avoid making subjective assessments of what should count as a novel word form (e.g. fucking vs. fucker). Further, homophones (words that sound the same but are spelled differently) were included (e.g., *phuck* versus *fuck*), as were hyphenated word combinations (e.g., *lullaby-lu* and *penis-breath*).

The only filter of the word lists was that each word was only counted once. By including a word once, this controlled for possible over-weighting of phonemes due to repetition of words in a chorus or due to common word usage. Note that this decision is conservative with respect to the hypothesis, as repeated use of a particular word might also be counted, arguably an indication of its special salience. Further, the conservative editing of the lists was done not only to avoid possible biased word selections but also to retain as many characteristics from the sources as possible. Without repeats removed, the sample involved 2398 words from lullabies, 2959 from carols, 3165 from heavy metal lyrics and 1084 from the swearword/profanity list for a total of 9606 words. With repeats removed the sample involved 682 words for lullabies, 829 for carols, 887 words for heavy metal lyrics and 437 for the swearword/profanity list for a total of 2835 words.

3.3 Phoneme Analysis

Words were formed into lists for each word source and were then phonemically transcribed, with a focus only on consonants (and not vowels), using a broad transcription process. Consonants were transcribed into the twenty-four English consonants (Ladefoged 2001).

The word lists so transcribed ultimately produced 8653 consonant phonemes for analysis of which 2340 were for carols, 1903 were for lullabies, 2663 were for heavy metal songs and 1747 were for swearwords/profanity. Consonants were characterized for their relative harshness or smoothness according to an independent sonority scale (Parker 2002, 2008). For a table of consonants organized in this way, see Table 3.1.

3.4 Results

Four sign tests were conducted comparing the carol list with swearword/profanity list, the carol list with the heavy metal list, the lullaby list with the swearword/profanity list, and the lullaby list with the heavy metal list. After organizing the consonants into a sonority hierarchy, the most sonorous consonants were compared against the least sonorous consonants. It was hypothesized that the positive affect inducing word sources (lullabies and carols) would contain relatively more sonorous consonants than the negative affect inducing word sources (swearword/profanity and heavy metal). Further, it was hypothesized that the negative

affect inducing word sources would contain relatively more harsh sounding consonants than the positive affect inducing word sources. Thus, for half of the sign test comparisons, one set of hypotheses are tested in one direction and for the other half of the sign test comparisons the hypotheses are tested in the other direction. For example, when comparing lullabies with heavy metal lyrics it was hypothesized that lullabies would have relatively more sonorant consonants but have relatively fewer plosive consonants.

When comparing the relative frequency of consonants for carols with the swearword/profanity list, there was a significant difference consistent with the hypothesis, with 19 of 24 cases showing differences in the predicted directions (19/24, p = 0.0033). However, when comparing carols with heavy metal there was not a significant difference (14/24, p = 0.2706). When comparing lullabies with the swearword/profanity list, there was a significant difference corresponding with our hypotheses (19/24, p = 0.0033). Again, however, when comparing lullabies with the heavy metal list there was no significant difference (13/24, p = 0.4194).

Due to the lack of clarity in terms of how to rank consonant classes near the center of the sonority hierarchy (Parker 2002), another sign test was conducted whereby only the top six most sonorous and six least sonorous consonants were used. Although fewer consonants are used in this analysis, the benefit is that theoretically troublesome consonant classes such as the voiced affricates and voiced fricatives are avoided. Further, this analysis also contains almost every consonant from Study 1 and

Study 2 with the exception of [n] and can thus be viewed as a narrower but also more refined test of previous findings.

When comparing the lullaby list with the swearword/profanity list, there was once again a significant difference that corresponded with our hypotheses (12/12, *p* = 0.0002). However, when comparing the lullaby list with the heavy metal list there was still not a significant difference (9/12, *p* = 0.0729). When comparing the relative frequency of consonants for the carol list with the swearword/profanity list, there was again a significant difference that corresponded with our hypotheses (12/12, *p* = 0.0002). However, when comparing the carol list with the heavy metal list, there was not a significant difference (7/12, *p* = 0.3872).

A series of chi-square analyses were also conducted using the same paired comparisons as the sign tests. The only difference is that instead of comparing percentages, the raw phoneme count data were used and collapsed such that the most sonorous phonemes are combined into one group and the least sonorous phonemes are combined into the other. This was done in one of two ways, in the overall chi-square analysis the top twelve most sonorous phonemes were compared against the bottom twelve. In the narrower analysis, the top six most sonorous consonants were collapsed and compared with the bottom six least sonorous consonants.

For the overall chi-square analysis, there were significant differences when comparing the lullaby list with the swearword/profanity list as well as a significant difference when comparing the carol list with swearword/profanity list, indicating that

the swearword/profanity had relatively more harsh consonants and fewer sonorant

consonants (χ^2(1, N = 3650) = 94.88, $p < 0.001$; χ^2(1, N = 4087) = 124.04, $p < 0.001$).

However, there were no significant differences when comparing the lullaby list with

heavy metal list nor when comparing the carol list with the heavy metal list (χ^2(1, N =

4087) = 0.00135, p = 0.97; χ^2(1, N = 5003) =1.32, p = 0.25).

For the chi-square analyses involving only the top six most sonorous and the

bottom six consonants, there were significant differences when comparing the lullaby

list with the swearword/profanity list as well as a significant difference when comparing

the carol list with the swearword/profanity list, indicating that again the

swearword/profanity list contained relatively more harsh consonants and fewer

sonorant consonants (χ^2(1, N = 2164) = 121.87, $p < 0.001$; χ^2(1, N = 2384) = 141.70, $p <$

0.001) . However, there were no significant differences when comparing the lullaby list

with the heavy metal list nor when comparing the carol list with the heavy metal list

(χ^2(1, N = 2546) = 0.86, p = 0.35; χ^2(1, N = 2766) =1.93, p = 0.16).

To test for differences within purported harsh and soothing word groups a series

of chi square tests were conducted. These chi square tests compared the carol list with

the lullaby list as well as the swearword/profanity list with the heavy metal list in a

general phoneme analysis (all 24 phonemes) and a narrower analysis (top 6 and bottom

6 phonemes). There was no significant difference between the carol and lullaby list for

both the overall analysis and narrower analysis (χ^2(1, N = 4243) = 1.03, p = 0.30; χ^2(1, N =

2360) =0.14, p = 0.70); however, there was a significant difference between the heavy

metal list and the swearword/profanity list for both the overall and narrower analysis with the swearword/profanity list containing more harsh consonants and fewer sonorant consonants than would be expected (χ^2(1, N = 4410) = 107.7, p <0.001; χ^2(1, N = 2570) =120.09, p < 0.001).

3.5 Discussion

The general findings of this study support the Affect Induction perspective in finding that words variously containing either harsh, plosive consonants or smooth, sonorant consonants are used differentially in real language contexts associated with variable affective tones. Thus, swearwords/profanity contained a higher proportion of words with harsh, plosive consonants than did lullabies, which contained a higher proportion of words with smooth, sonorant consonants. For a bar graph depicting the proportion of least sonorous versus most sonorous consonants for each word source, see Figure 3.1. Part of the reason the swearword/profanity word list contained so many instances of harsh, plosive consonants is because some words or word combinations appeared many times (e.g., *shit* and *fuck*). For example, *fuck* was found in 64 words and *shit* was found in 29 words. Although it could be argued that these words are biasing the results, it could also be argued that that bias is in fact simply reinforcing the affective effects of these words.

3.5.1 Swearwords and Profanity

The clearest statistical effects obtained in this analysis were for swearwords/profanity where the consonant associations were significant in every test. It is possible that swearwords represent a special class of words because they are so singularly emotionally charged and are used in situations of conflict to intimidate or bluff. Swearing may even activate a preparatory "fight-or-flight" response for such situations of aggression (Stephens, Atkins & Kingston 2009). For example, swearing appears to have the effect of increasing heart rate and also reduces pain. Stephens and colleagues (2009) had participants hold their hand in ice water to compare their perception of pain and how long they were able to hold their hand under water when they swore versus when they said control words. Participants were able to hold their hand under water for significantly longer periods of time while swearing and perceived less pain than when they only said control words (Stephens et al. 2009).

Swearwords and swearing may also be distinctive neurologically in being processed on a lower (limbic) level than words that are primarily referential, and thus cortically processed (Lancker & Cummings 1999). For example, some individuals with various neuropathologies such as Tourette syndrome as well as individuals with language aphasias are known to vocalize epithets and profanities (Lancker & Cummings 1999). Some of these aphasic individuals readily exhibit interjections of a vulgar nature, despite not being able to say anything else (Lancker & Cummings 1999). Swearing is considered part of "automatic" speech and even individuals with aphasia that struggle

to say other words are able to swear fluently with appropriate prosody. After analyzing

the phonemic content of words related to these outbursts, Lancker and Cummings

(1999) found that words containing the phonemes /F/, /SH/ and /K/ were the most

common; though, the authors offer no suggestion as to why this is the case.

Lancker and Cummings (1999) argue that the limbic system and basal ganglia are

critically involved for these involuntary epithets due to their role in emotional

processing, motivation and social behaviour. Further, the limbic system has also been

found to be involved with animal calls. For example, stimulation of various areas of the

limbic system in macaques and squirrel monkeys induces the animal to produce various

emotionally charged vocalizations (Robinson 1967, Jurgens & Ploog 1970). Some of the

stimulation induced macaque calls were described as loud, plosive and harsh as well as

shrill and guttural (Robinson 1967), adjectives that can also describe swearing in

contexts of aggression. Indeed, pejorative epithets may be the closest analogue to

agonistic calls humans have, aside from actually growling or screaming.

3.5.2 Heavy Metal

Heavy metal was chosen as a candidate word source for increased use of harsh

consonants because of the negative affect inducing effects that are presumed to be

associated with this genre of music. However, heavy metal lyrics proved enigmatic in

the current analysis, as this was the only word source that did not yield consonant-type

effects consistent with my predictions. Part of the reason for these results may be that

much of the acoustic impact from heavy metal songs is produced not by the vocals but by the instruments such as the drums and the electric guitar. In addition, the vocal style of heavy metal singers may exert a greater influence than the lyrics themselves: for example, screaming, shouting and growling are used in some heavy metal sub genres such as death metal. Hence, the negative affective tone associated with heavy metal may have more to do with instrumental harshness and harshness of voice quality than with the kind of inherent harshness of particular consonants represented in lyrics. Furthermore, it is also the case that not all heavy metal songs are so obviously aggressive in tone. Indeed, there can be considerable variation in the content of mood of heavy metal songs, with many examples of melancholic songs or love ballads.

An additional limitation is that this analysis focused on written words rather than the actual acoustic signal. Properties such as intonation, accent and stress are not accounted for in a written analysis. Perhaps a more pertinent test of the Affect Induction model would be to analyze lyrics as they are actually produced acoustically in each of the genres tested. The Affect Induction model would predict not only that the consonant content of words in these genres might differ consistently but how the words were actually produced would also differ in order to increase or emphasize the relevant affective distinctions.

3.5.3 Parker's Sonority Hierarchy

This survey research was analyzed using Parker's sonority hierarchy (2008). It should be noted, however, that sonority itself is a debated topic within linguistics with some question even about its reality (Ohala 1990). Part of the issue is that there is no agreed upon method for measuring sonority, nor has there been a clear theoretical rationale for why it might exists (Clements 1990), except that the Affect Induction perspective now offers such an account grounded in animal biology and the evolution of signaling broadly (see below).

Historical ambiguity notwithstanding, there have been attempts at defining and quantifying sonority, many of which are language specific. Among the most extensive investigations are those by Parker (2002, 2008) who has attempted to categorize more phonemes than any previous scheme. Parker (2008) has also collected data not only from English speakers but also speakers of Spanish and Quechua and his results demonstrate a cross-linguistic pattern supporting his sonority hierarchy. For these reasons Parker's sonority hierarchy was used. Parker's work investigated the sonority hierarchy by having participants speak specific words while recording various aspects of articulation such as interoral air pressure, the intensity, frequency of the first formant, total segmental duration, peak intraoral air pressure, and combined oral plus nasal air flow. The measure that consistently correlated best with sonority was intensity or amplitude (loudness) with a very strong Spearman correlation of 0.91 (Parker 2002, Parker 2008).

Although Parker's sonority scale (2008) focuses on variable intensity, it was a useful tool for categorizing phonemes on a harshness-smoothness scale and it is likely that this reflects natural covariation in phoneme intensity and harshness. Hence, the Affect Induction model might help to provide the previously missing theoretical rationale for the psychoacoustic aspects of the sonority scale. Consonants low in sonority tend to have a broadband harsh pattern; conversely consonants high in sonority tend to have tonal, soothing pattern. This difference in acoustic characteristics may also account for findings on the "agreeableness" and "disagreeableness" of consonants (Roblee & Washburn 1912). For example, the consonants /G/, /K/ and /T/ in vowel-consonant nonsense words were deemed the most "disagreeable" and /L/, /M/, and /N/ were deemed the most "agreeable" by participants (Roblee & Washburn 1912).

Overall, the findings from this sampling of real word corpuses suggest that words used in different contexts might be chosen based in part on the affective impact of their phonemic content as predicted by the Affect Induction approach to animal communication (Owren & Rendall 1997). Swearwords/profanity in particular contain more harsh sounding consonants and fewer sonorant sounding consonants relative to lullabies and carols. A corollary of these findings is that the consistent word-form expectations demonstrated by experimental participants in my earlier experiments, and in other similar experimental studies of the bouba-kiki phenomenon, might not be limited to nonsense images and words but ultimately affect real word usage, at least in some natural language contexts. This outcome runs counter to the Saussurean linguistic

dictum that the connection between word form and referent is strictly arbitrary and

conventional.

Sonority Hierarchy

Natural Class	Index	Phonemes in analysis
glides	12	([j]*y*) [w]
rhotic approximants (ɹ)	11	([ɹ]*r*)
flaps	10	
laterals	9	[l]
trills	8	
nasals	7	[n] [m] ([ŋ]*ing*)
voiced fricatives	6	[v] [z] ([ð]*th*at)
voiced affricates	5	([dʒ] *j*) ([ʒ]trea*s*ure)
voiced stops	4	[b] [d] [g]
voiceless fricatives (including h)	3	[f] [s] ([θ]*th*ink) [h]
voiceless affricates	2	([ʃ]*sh*) ([tʃ]*ch*)
voiceless stops (including ʔ)	1	[k] [t] [p]

Table 3.1 A table of the twenty-four English consonants organized into a sonority

hierarchy by word class (Parker 2008). Note that English consonants do not typically

contain trills (rolling of 'r's) or flaps.

Relative Percentage of Consonants Based on Sonority
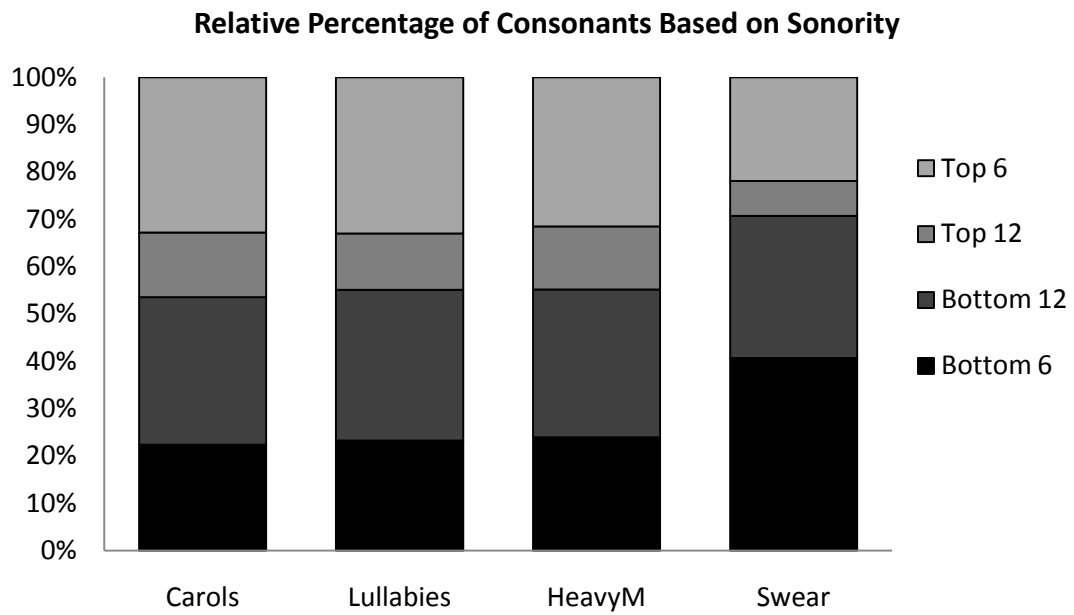
Figure 3.1. A bar graph depicting the relative percentage of the top twelve consonants in terms of sonority versus the bottom twelve consonants. The relative percentage of the top 6 consonants in terms sonority and the bottom 6 consonants are also included.

CHAPTER FOUR

The purpose of this thesis was to investigate what implications the Affect Induction model of animal communication might have on human communication. Studies 1a, 1b and 2 were conducted on the bouba-kiki phenomenon in order to test the Affect Induction model in a presumably arbitrary context. The nonsense words and randomly generated images used in these studies were methodologically beneficial because they controlled for the confounds of actual language use and associations with real world objects; however, this artificial aspect also makes the results from these studies less applicable to real world language use. Thus, an analysis was conducted on the phoneme content of real words from various sources presumed to induce different emotional responses. The results from these studies generally supported the idea that the Affect Induction model might have some natural application to human communication as well, including language.

4.1 The Non-Arbitrariness of Words

Despite the nonsense nature of the words and randomness of the images used, participants' responses in study 1a, 1b and study 2 were not insensitive to the sonority of the consonants. If words are arbitrary and how they sound has no inherent meaning then a random response would be expected. Instead, participants in study 1a and 1b more commonly selected nonsense words with plosive consonants for more jagged images and more commonly selected words with sonorant consonants for rounded

images. Further, participants in study 2 manipulated the image to be more jagged for words with plosive consonants and to be smoother for words with sonorant consonants. Both of these studies strongly suggest a non-arbitrary relationship between the acoustic properties of the nonsense words and the visual properties of the images. This controlled investigation into the bouba-kiki phenomenon suggests that the types of consonants used have a significant impact on participant responses.

Based on the corpuses of words used for analysis in chapter 3, real words may also be affected by how their acoustic properties impact listeners. The swearwords and profanity category contained significantly more harsh sounding consonants than lullabies and carols. However, the heavy metal category results did not coincide with our hypotheses as they did not significantly differ from lullabies or carols but were significantly different from the swearword/profanity category. Overall, the swearword/profanity category had the strongest a priori prediction based on the Affect Induction model due to agonistic contexts in which expletives and epithets are sometimes used. This finding undermines the linguistic tenet of arbitrariness because if words were truly arbitrary then it would be expected that there would have been no observed significant difference among the word samples in terms of consonant content proportions. Taken together, these studies provide evidence that the Affect Induction model may have explanatory power not only for animal communication but also for human communication.

4.2 Implications and Future Research

The most straightforward implication of this thesis is that language is not completely arbitrary. If words were completely arbitrary and had no inherent connection with their referents, then we would expect no basis by which participants could match words with images in the bouba-kiki studies nor would there be relative phoneme differences among the real corpuses studied. Any observed non-random response tendency from participants and any relative phoneme differences among word sources is enough to undermine the dictum of arbitrariness; however, these findings may have broader implications because the direction of results were specifically predicted by the Affect Induction model.

The Affect Induction model posits that animal communication is about signalers producing sounds with acoustic properties that best influence or manipulate a targeted listener's affective state and not about relaying information from one animal to another, as purported by classical ethology (Owren & Rendall 1997). From an evolutionary perspective, giving away information freely is altruistic and is not an evolutionarily stable strategy as any selfish individual can take advantage of the system without contributing (Krebs & Dawkins 1978). Just as classical ethologists viewed animal communication as an exchange of information among senders and receivers, linguists have traditionally viewed human language as being exclusively purposed to convey information to other individuals (Scott-Phillips 2006). From this free-exchange of

information perspective, selection has been on language to improve this conveyance of information by increasing understandability (Pinker & Bloom 1990).

The alternative to viewing language as altruistic is to view it as selfish: it benefits the signaler or speaker directly. Perhaps this direct pay-off has been the main selection pressure for its evolutionary development (Scott-Phillips 2006). Speculatively, it could have been the case that early in language evolution, perhaps even before human language, the types of sounds used could have conveyed meaning via the sound patterns themselves; however, as language became more abstract and complex this lower level aspect became less readily apparent and now plays a more subtle role in modern discourse and language use.

4.2.1 Language Learning

Although sound symbolism may only play a subtle role in modern language, it might be crucially important for language learners. One critical issue in forming a symbol system is the grounding of that system with meaning; otherwise an infinite regress can occur (Hanard 1990). In other words, without a starting point, it is impossible for a symbol system to represent anything meaningful as it would require using other symbols that are also devoid of meaning. Sound symbolism offers a way out of this problem by grounding the symbol system in inherent sound-meaning relationships. Inherent sound symbolic associations can then bootstrap the symbol system by layering further meanings and complexities on this initial base layer (Cowley 2007). Cowley

(2007) suggests that innate biases and affect-based dynamics between young infants and caregivers is how the bootstrapping of language learning is initiated. An important innate bias may well be a simple distinction between abrasive acoustic patterns (shhh!) and harmonious acoustic patterns (infant directed speech) with negative and positive affect, respectively. A simple example of this affect-based dynamic is the affect communicated by infant directed speech and the response of the infant toward the caregiver through crying and coos. Language learning is further built upon by learning more complex embodied "body-world" coordination and eventually toward using utterances of words to alter the behaviour of caregivers and others (Cowley 2007).

The fact that researchers have found that 2.5 year old children respond non-randomly to a bouba-kiki task suggests that there are inherent sound symbolic associations (Maurer et al. 2006). Although Maurer argues that her findings are due to the vowel content of the words, the findings of this thesis suggest that the acoustic properties of consonants are also important in driving the effect. More work with pre-linguistic infants is necessary to further investigate the exact nature of these sound symbolic associations and to see if they are somehow detectable even before language learning and usage. Though it may be difficult to study infants independent from the effects of learning as it may be the case that as soon as they are born they are exposed to language use enough to form statistical learning associations.

If sound symbolic symbol grounding is correct, then presumably words that suit their referent will be more readily learned. There is some evidence that supports this

conclusion; for example, researchers found that 3 year old Japanese children were able to generalize action words, or verbs, significantly more if they were given mimetic words that sound symbolically mimicked the action, as opposed to words that did not (Imai, Kita, Nagumo, Okada 2008). Verbs are particularly difficult for children of this age to learn, yet in this study they were able to successfully generalize observed actions over 80% of the time when given sound symbolic words but responded at chance levels when given words without the sound symbolic aspect (Imai et al. 2008). Based on related research, Japanese mothers preferentially used mimetic sound symbolism 57% of the time when speaking to their children about a particular action and used these sound symbolic words only 12% of the time when describing the same actions to a researcher (Nagumo, Imai, Kita, Haryu, & Kajikawa 2006). This finding suggests that mothers intuitively used the sound symbolic words, as opposed to conventional words that lacked sound symbolism, in order to help their children learn.

There is further evidence that sound symbolism can play an important role in language learning, even for older children and adults (Parault & Parkinson 2008, Parault & Schwanenflugel 2006, Parault 2006). For example, 5th and 6th graders were better able to guess the definitions of words that were sound symbolic and demonstrated improved word learning for sound symbolic words versus words that were not sound symbolic (Parault & Parkinson 2008). Further, adults were also better able to recognize sound symbolic obsolete words versus obsolete words that were not sound symbolic and they were able to generate better definitions for sound symbolic words in contrast to words that were not sound symbolic (Parault & Schwanenflugel 2006).

4.2.2 Language Change

Language learning is sometimes characterized as being a one way street whereby the language community impinges itself upon the language learner; however, the language learner could have considerable impact on the language community (Clark & Roberts 1993). After all, languages are constantly in flux, with words and phrases going out of use and new words coming into use. Human beings live in generational cohorts with the baton of language being passed on to the next, sometimes with changes along the way (Clark & Roberts 1993). Words that are more easily learned and retained because they are sound symbolically suited to their referent could very well replace less suitable synonyms and variants of that word. Further, words that are more effective in inducing affect may be particularly pressured to be retained. One example of how sound symbolism can influence word survival is the cross-cultural similarities for kin terms, particularly for parents. For example, words designating mothers contain nasals (e.g. /M/, /N/) and words designating fathers contain plosives (e.g. /T/, /P/) cross-culturally more often than would be expected by chance (Murdock 1959). Murdock (1959) found that of the 474 languages studied 52% of words for mothers contained the nasals /M/ and /N/ while only 15% of words for fathers contained these consonants; conversely, 56% of words for fathers contained the plosives /T/ and /P/ while only 7% of words for mothers contained these consonants. Of course not all words are sound symbolic as there are far too many words, each requiring a distinct sound pattern;

however, sound symbolism may still play an ongoing subtle role in word survival

(Jespersen 1922).

Studying language as an arbitrary information exchange system ignores or at

least trivializes the importance of the physical acoustic reality of language. Sound itself

is a traveling oscillating wave of pressure changes in a medium, typically air, often

eventually impinging itself on an auditory system. The characteristics of speech sounds

are critically important for differentiating one word from another and the characteristics

of speech sounds are critical for sound symbolism. One way of conceptualizing the

acoustic differences in the consonants studied in this thesis is to consider sonority as a

scale shifting from harsh abrupt consonants towards increasingly vowel-like consonants

(Jakobson & Waugh 1979). For example, the stop consonants /p/, /t/, /k/ are

characterized phonemically as being the result of closure of the vocal tract with a jut of

air forced through at release of the closure resulting in a noisy burst of sound

(Ladefoged 2001). This noisy burst of sound is visible in spectrograms as a short chaotic

broadband pattern. In contrast, the nasal consonants /m/, /n/, /ŋ/, are voiced

throughout, meaning the vocal folds are vibrated throughout articulation. During

articulation of nasal consonants, sound is released up through the nasal cavity; resulting

in a sound pattern much like vowels and is visible in spectrograms as having distinct

formant frequencies (Ladefoged 2001). The distinction between these consonants is

that the stop consonants appear much more chaotic than the ordered and harmonic

nasal consonants. This distinction has also been characterized as a difference between

"strident" consonants that restrict air-flow and approximate noise, versus "mellow"

consonants that permit air flow and approximate vowels (Jakobson & Waugh 1979).

Voicing, in general, appears to be important for Parker's sonority hierarchy as voiceless

plosives, voiceless affricates and voiceless fricatives are all lower in sonority than their

voiced counter parts, adding support to the conceptualization of the sonority scale as

increased tendencies towards being vowel-like. Instead of simple arbitrary units of

words, phonemes and their articulation have real world physical differences in sound

patterning that potentially impact the listener in psychologically different ways.

Sound symbolism in general may be an example of a cross-linguistic universal.

Size-sound symbolism has been reported cross-linguistically and Parker's (2008) sonority

hierarchy is applicable to Spanish and Quechua as well as English, thus suggesting that

there is a universal tendency toward a sonority hierarchy; though, like other examples

of sound symbolism, the actual speech sounds of specific languages will augment and

alter the specifics of the relationships (Parker 2002). The Affect Induction model of

communication may also prove to be cross-linguistically universal and may even be an

example of a cross-species universal, at least among vertebrates. Bird species as well as

non-human primate species have been found to produce harsh grating calls in agonistic

contexts and more harmonious calls in social contexts (Morton 1977). This cross species

use of affect inducing sound patterns may be due to shared auditory systems and brain

areas in the mid-brain that are responsible for emotional responses. Central auditory

brain areas, such as the primary auditory nucleus and auditory midbrain, have remained

relatively conserved across vertebrate species; this suggests that there could be shared

processing of acoustic information (Moss & Carr 2003). Even some fish, such as

mormyrids, have neurons in the central nucleus that are sensitive to the growls and moans of their conspecifics (Crawford 1997). Further, it also appears that some species of fish also use affect induction similarly to their terrestrial vertebrate relatives. For example, the male midshipman fish (*Porichthys notatus*) produces harmonic hums to attract females and grunts that are broadband signals to conspecifics in agonistic situations (Bass, Bodnar & Marchaterre 1999).

One behavioural response to a sound that coincides with the Affect Induction model is the acoustic startle response (Koch 1999). The acoustic startle response is a reaction to a sudden aversive sound resulting in tensing up and often retracting away from the sound source, presumably to avoid injury or death. This response has its neurological roots in the lower brain stem and is suggested to be homologous with the fish flight response, also mediated by the brain stem (Koch 1999). This obligatory auditory pathway may play a low level role even for language sound patterns.

4.2.3 The Importance of Affect

The fact that words are representational obscures the importance of the sound patterns and the contexts in which they are used. In certain contexts the emotional impact comes not from the representational aspect of the words said but their acoustic aspects and how they were pronounced. Certainly in the case of swearwords the emotional impact of the words take precedence over their representational meanings. An anecdotal every day example of how the affective impact of words can sometimes be more relevant than the representational meaning is when an exchange of words results

in offence and someone saying "it is not *what* you said but *how* you said it". All words may carry a certain connotative or emotional meaning that "piggy-backs" on their representational meaning and verbal usage (Besnier 1990).

Osgood and colleagues (1957) attempted to measure the meaning of words by having participants rate words based on a number of different scales. This analysis was called a "semantic differential" and measured the "semantic space" of words. Analyzing the semantic differential involved conducting a principle component analysis on these response scales. Osgood and colleagues (1957) found that three factors contributed most to the word ratings; these three factors were an evaluative factor (good-bad), a potency factor (weak-strong) and an activity factor (fast-slow) (Osgood 1957: 44-46). The evaluative factor was particularly important as it accounted for 50-75% of the variance across analyses. The evaluative factor included ratings along good-bad, optimistic-pessimistic, complete-incomplete scales. This measurement of "semantic space" primarily accounts for the affective or emotional aspect of words; however, it does not measure the representational meaning of words. It is likely that words convey both a higher level representational aspect while simultaneously conveying affective properties that affect what we say and how we say it; particularly when speaking in contexts involving social interaction with other individuals in varying contexts of negative and positive social affiliation.

4.2.4 Animal Research

One important avenue of research for testing the Affect Induction model is animal research. For example, a primate species could be used to see if abrasive sounds increase physiological arousal and sonorant sounds reduce arousal. Similarly, conditioning studies can be conducted to see if animals more readily form an association between a harsh sound with an electric shock versus a sonorant sound with an electric shock; or conversely, if positive reinforcement is more readily associated with a sonorant sound versus an abrasive sound. Almost any sound can be used to form conditioned responses, however, the number of trials or time it takes to form these associations could be measured in order to test the Affect Induction model. Further, the retention of these conditioned associations could be tested with the hypothesis that the harsh sound will be associated with the noxious stimuli or electric shock much longer and will take longer to extinguish than the sonorant sound with noxious stimuli association.

There have already been studies that examine how different sound patterns affect animals, especially as a means for pest control (Bomford & O'Brien 1990). Researchers have found that broadband "noisy" sonic deterrents are the most effective at dispersing pest animals; however, these effects are temporary as animals eventually habituate to these sounds (Bomford & O'Brien 1990). In contrast, high frequency tonal pest control sonic devices appear to have no effect at all. These results are exactly what would be predicted based on the Affect Induction model. Broadband harsh calls that are species specific appear to be particularly effective and resistant to habituation, probably

because species specific calls also coincide with other agonistic behaviours such as physical injury (Bomford & O'Brien 1990). For example, starlings (*sturnus vulgaris*) were presented three types of sounds: a pure tone sound, a broadband white noise "hissing" sound and a starling distress call (Johnson, Cole & Stroup 1985). The pure tone sound had little to no effect on the starlings; however the white noise and starling distress calls both instilled a fear response, though the distress call was much more resistant to habituation (Johnson et al. 1985).

Squirrel monkeys that were socially isolated after birth and raised in a cage with a surrogate cloth "mother" also exhibit behavioural responses to sounds that coincide with the Affect Induction model (Herzog & Hopf 1983). A novel test object was placed into the cage of the monkeys and it produced one of several different sounds through a speaker. Most sounds were species specific, including the socially relaxed "twitter" and "play peep" sounds as well as the moderately aversive "cackle" and slightly aversive "caw" sounds (Herzog & Hopf 1983). The aversive "cackle" and "caw" sounds led the monkeys to cling to their surrogate cloth mother and remain in this defensive position, even after repeated sessions of these sounds. The twitter and play peep sounds initially resulted in the monkeys clinging to their surrogate cloth mother; however, these sounds were habituated to and had no subsequent effect after a few play-backs (Herzog & Hopf 1983). These findings are what would be expected based on the Affect Induction model, despite the fact that the most aversive species specific sounds such as shrieks and screams were not included in the test phase of this study because during a pilot study these sounds had a very strong inhibiting effect (Herzog & Hopf 1983).

Due to the potential cross-species universality of affect induction, many different test species may be used to investigate this hypothesis. For example, chinchillas would make a particularly interesting test species because they are able to differentiate between subtle differences in consonant phonemes and would thus be a good test species to see if less sonorous consonants are more readily associated with noxious stimuli versus more sonorous consonants and vice versa (Kuhl & Miller 1975). Additionally, chinchillas are known to produce broadband "noisy" barks in agonistic contexts and to produce tonal "gentle but, bright" contact calls in more social contexts and this coincides exactly with the Affect Induction model (Hunyady 2008).

4.3 Conclusion

The results presented in this thesis undermine the linguistic tenet of absolute arbitrariness of words with regard to their referent. The bouba-kiki phenomenon was investigated using a novel image generation method and by controlling consonant content and the results confirm a priori expectations based on the Affect Induction model. The phoneme frequency analysis demonstrated relatively more harsh sounding consonants than comparison groups for the swearword/profanity word group, also confirming a priori expectations based on the Affect Induction model. Further research is necessary to detect the nuances and reveal the full explanatory impact the Affect Induction model has on human communication. Ultimately, the Affect Induction model of communication may prove to be an underlying universal that is pervasive across languages and even across species.

REFERENCES

Aronoff, J. (2006) How we recognize angry and happy emotions in people, places and things. *Cross-Cultural Research*, *40*, 83-105.

Aronoff, J., Barclay, A. M., & Stevenson, L. A. (1988). The recognition of threatening facial stimuli. *Journal of Personality and Social Psychology*, *54*, 647-655.

Aronoff, J., Woike, B. A., & Hyman, L. M. (1992). Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion recognition. *Journal of Personality and Social Psychology*, *62*, 1050-1066.

Back, E. (2008). Retrieved from http://wordpress-plugins.feifei.us/politeifier/.

Bar, M. & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science*, *17*, 645-648.

Bar, M. & Neta, M. (2007). Visual elements of subjective preference modulate amygdale activation. *Neuropsychologia*, *45*, 2191-2200.

Bass A.H., Bodnar D.A., and Marchaterre M.A. (1999). Complementary explanations for existing phenotypes in an acoustic communication system. In Hauser M and Konishi M (eds) *Neural Mechanisms of Communication.* MIT Press, Cambridge, pp. 493–514.

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

Birkbeck, N. (2008). Retrieved from http://www.neilbirkbeck.com/jshape/single.html.

Bolinger D. (1964). Intonation as a universal. In Proc. 9th Int. Congr. Linguists, 1962, ed.
HG Lunt, pp. 833-44. The Hague: Mouton.

Bolinger D. (1978). Intonation across languages. See Greenberg et al 1978, pp. 471-524

Bomford, M. & O'Brien, P.H. (1990). Sonic deterrents in animal damage control: A
review of device tests and effectiveness. *Wildlife Society Bulletin, 18*, 411-422.

Brown R.W. (1958). *Words and Things*. Glencoe, IL: Free Press.

Brown R.W., Black A.H., Horowitz A.E. (1955). Phonetic symbolism in natural languages.
Journal of Abnormal Social Psychology, 50, 388-93.

Clark, R. & Roberts I. (1993). A computational model of language learnability and
language change. *Linguistic Inquiry, 24*, 299-345.

Clements, G. N. (1990). The role of sonority in core syllabification. In J. Kingston & M. E.
Beckman (Eds.), Papers in laboratory phonology I. Between the grammar and
physics of speech (pp. 283–333). Cambridge: CUP.

Cowley, S.J. (2007). How Human infants deal with symbol grounding. *Interaction Studies*,
8, 81-204.

Crawford, J.D. (1997). Feature-detecting auditory neurons in the brain of a sound-
producing fish. *Journal of Comparative Physiology*, *180*, 439-450.

Davies, N.B., Kilner, R.M. & Noble, D.G. (1998). Nestling cuckoos *Cuculus canorus*
exploits hosts with begging calls that mimic a brood. *Proceedings of the Royal
Society of London*, *265*, 673-678.

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

Dawkins, R. & Krebs, J.R. (1978). "Animal signals: information or manipulation?". *Behavioural Ecology: An Evolutionary Approach*. Oxford: Blackwell Scientific Publications. pp. 282-309.

Dawkins, R. & Krebs J.R. (1979). "Arms races between and within species". *Proceedings of the Royal Society of London. Series B, Biological Science ,205*, 489-511.

Fenton, M.B. & Licht, L.E. (1990). Why rattle snake? *Society for the Study of Amphibians*, *24*, 274-279.

Fitch, W.T. & Giedd, J., (1999). Morphology and development of the human vocal tract. A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*, 1511–1522.

Fitch, W.T., & Reby D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *268*, 1669-1675.

Gebels, G. (1969). An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior*, *8*, 310-12.

Haahr M. (2010). random.org - true random number service. Retrieved from http://www.random.org/.

Hanard. (1990). The symbol grounding problem. *Phsyica D 42*, 335-346.

Heavy metal. (n.d.). *The American Heritage® Dictionary of the English Language, Fourth Edition*. Retrieved July 16, 2010, from Dictionary.com website:http://dictionary.reference.com/browse/heavy metal.

Herzog, M. & Hopf S. (1983). Effects of species-specific vocalizations on behavior of surrogate-reared Squirrel monkeys. *Behaviour, 86*, 197-214.

Hunyady, H. (2008). Vocal sounds of the chinchilla (masters thesis). Retrieved from: http://etd.ohiolink.edu/send-pdf.cgi/Hunyady%20Heather.pdf?acc_num=bgsu1206318183.

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

Jakobson, R. & Waugh, L. (1979). *The sound shape of language*. Bloomington: Indiana University Press.

Jay, T. & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research, 4,* 267-288.

Jespersen, Otto (1922). *Language: Its Nature, Development and Origin*, London: Allen & Unwin.

Jespersen, O. (1933). "Symbolic value of the vowel *i*," In O. Jespersen*, Linguistica: Selected papers in English, French and German*. Copenhagen: Levin and Munksgaard. pp. 283-303.

Johnson, R., Cole, P. & Stroup, W. (1985). Starling response to three auditory stimuli. *Journal of Wildlife Management*, *49*, 620-625.

Jones R. (2010). Retrieved from www.noswearing.com/dictionary

Jurgens, U. & Ploog, D. (1970). Cerebral representation of vocalizations in the squirrel

monkey. *Experimental Brain Research*, *10*, 532-554.

Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism.

Marketing Letters, *11*, 5-20.

Klink, R.R. (2003). Creating meaningful brand names: The relationship between the

brand name and brand mark. *Marketing Letters, 14*, 143-157.

Kluytmans T. (2010). Kididdles the best in children's music. Retrieved from

http://www.kididdles.com/lyrics/lullabies.html

Köhler, W. (1947). *Gestalt psychology* (2nd edn.). New York: Liverright Publishing.

Koch, M., (1999). The neurobiology of startle. *Progress in Neurobiology*. 59, 107-128.

Kuhl, P., & Miller, J. (1975). Speech perception by the chinchilla: Voiced-voiceless

distinction in alveolar plosive consonants, *Science*, *190*, 69-72.

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sound of*

Language. Los Angeles: Blackwell Publishing.

Lancker, D.V. & Cummings, J.L. (1999). Expletives: Neurolinguistics and neurobehavioral

perspectives on swearing. *Brain Research Reviews*, *31*, 83-104.

Lapolla, R.J. (1994). Some observations on the function of sound in clinical work. In:

Hinton L., Nichols, J. & Ohala, J.J. *Sound Symbolism.* New York: Cambridge

University Press.

Lowrey, T.M. & Shrum, L.J. (2007). Phonetic symbolism and brand name preference. *Journal of Consumer Research*, *34*, 406-414.

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

Maxwell, D. (2010). Retrieved from http://current.com/entertainment/comedy/90396735_227-words-you-cant-say-on-youtube.html.

McDonnell, R. (2010). Retrieved from http://www.darklyrics.com.

Miron, M.S. (1961). A cross-linguistic investigation of phonetic symbolism. *Journal of Abnormal Social Psychology, 62*, 623-30.

Morton, E.W. (1977). "On the occurrence and significance of motivation-structural rules in some bird and mammal sounds," *American Naturalist, 111*, 855-869.

Moss, C.F. & Carr, C.E. (2003). Comparative psychology of audition. In: *Handbook of Psychology*. Gallagher, M., Nelson, R.J. Nelson, Weiner, I.B.(Eds). Hoboken, New Jersey: John Wiley & Sons.

Morton, E.S. (1994). Sound symbolism and its role in non-human vertebrate communication. In: Hinton L., Nichols, J. & Ohala, J.J. *Sound Symbolism.* New York: Cambridge University Press.

Murdock, G.P. (1959). Cross-language parallels in parental kin terms. *Anthropological Linguistics*, *1*, 1-5.

Nagumo, M., Imai, M., Kita, S., Haryu, E., & Kajikawa, S. (2006). Sound iconicity

   bootstraps verb meaning acquisition. Paper presented at the *XVth International*

   Conference of Infant Studies, Kyoto, Japan.

Negus, V.E. (1949). *The comparative anatomy and physiology of the larynx*. New York:

   Hafner Publishing Company.

Newman, S.S. (1933). Further experiments in phonetic symbolism. American Journal of

   Psychology, *45*, 53-75.

Nuckolls, J.B. (1999). The case for sound symbolism. *Annual Review of Anthropology*, *18*,

   225-252.

Ohala, J.J. (1996). Speech perception is about hearing sounds, not tongues. *Acoustic*

   Society of America, 99, 1718-1725.

Ohala, J.J. (1990). There is no interface between phonology and phonetics: a

   personal view. *Journal of Phonetics, 18*, 153-71.

Ohala, J.J. (1994).The frequency code underlies sound symbolic use of voice pitch. In:

   Hinton L., Nichols, J. & Ohala, J.J. *Sound Symbolism.* New York: Cambridge

   University Press.

Ohala, J.J. (1996). Speech perception is hearing sounds, not tongues. *The Journal of the*

   Acoustical Society of America, *99*, 1718-1725.

Ostwald, P.F. (1994). Some observations on the function of sound in clinical work. In:

   Hinton L., Nichols, J. & Ohala, J.J. *Sound Symbolism.* New York: Cambridge

   University Press.

Osgood,C. E., Suci,G. J.,&Tannenbaum,P. H. (1957).*The measurement of meaning*. Urbana: University of Illinois Press.

Ottoson, U. Bäckman, J. & Smith, H.G. (1997). Begging affects parental effort in the pied flycatcher, *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology, 41*, 381-384.

Owren, M.J., Rendall, D. (2001) Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, *10*, 58-71.

Owren, M.J. & Rendall , D. (1997). An affect-conditioning model of nonhuman primate vocal signaling. In *Perspectives in Ethology* (Vol. 12). Owings et al (eds). New York: Plenum Press.

Owren, M.J. & Rendall, D. & Ryan M.J. (2010). Redefining animal signaling: Influence versus information in communication. *Biology and Philosophy*, DOI 10.1007/s10539-010-9224-4.

Parault, S. J. (2006). Sound symbolic word learning in context. *Contemporary Educational Psychology, 31*, 228-252.

Parault, S. J., & Parkinson, M. (2008). Sound symbolic word learning. *Contemporary Educational Psychology*, doi:10.1016/j.cedpsych.2007.07.003.

Parault, S. J., & Schwanenflugel, P. J. (2006). Sound symbolism: A possible piece in the puzzle of word learning. *Journal of Psycholinguistic Research*, *35*, 329-351.

Parker, S. (2002). *Quantifying the sonority hierarchy*. Ph.D. thesis: University of Massachusetts, Amherst.

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, *36*, 55-90.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Brain and Behavioral Sciences, 13*, 707-727.

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

Reilly, J., Biun, D., Cowles, H.W., & Peelle, J.E. (2008) Where did words come from? A linking theory of sound symbolism and natural language evolution. *Nature Precedings*. Retrieved from:

http://precedings.nature.com/documents/2369/version/1.

Rendall, D., Notman, H. & Owren, M.J. (2009) Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *Journal of the Acoustical Society of America*. 125: 1792-1805.

Rendall, D., Owren, M.J. & Ryan, M.J. (2009) What do animal signals mean? *Animal Behaviour, 78*, 233-240.

Rendall, D. & Owren, M.J. (2009) Vocalizations as tools for influencing the affect and behavior of others. In: S. Brudzynski (Ed.) *Handbook of Mammalian Vocalizations*. Oxford: Academic Press, pp. 177-186.

Robinson, B.W. (1967). Vocalizations evoked from forebrain in Macaca mullata. *Physiology and Behavior*, *2*, 345-354.

Roblee, L. & Washburn, M.F. (1912). The affective values of articulate sounds. *The American Journal of Psychology*, *23*, 579-583.

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

Saussure F de. (1959). *Course in General Linguistics*, Bally, E.C. Sechehaye, A. Riedlinger, A. (eds) Transl. W Baskin. New York: McGraw-Hill.

Smith, J. (2010). Retrieved from http://www.christmas-carols.net/.

Scott-Phillips, T. C. (2006) Why talk? Speaking as selfish behaviour. In *Proceedings of the 6th International Conference on the Evolution of Language*, pp 299-300.

Stephens, R., Atkins, J., & Kingston, A. (2009). Swearing as a response to pain. *NeuroReport, 20*, 1056-1060.

Thompson, L. (2010). Retrieved from www.jomsocial.com/download/addons/content-enhancements/profanity-filter-word-list.html.

Ultan, R. (1978). Size-Sound Symbolism. In: Universals of human language (Vol. 1), Greenberg, J.H. (Ed). pp. 525-528.

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

Wilson, T.V. Retrieved from: http://science.howstuffworks.com/lrad.htm (Accessed, February 18th 2010).

Wise, K.K., Connover, M.R. & Knowlton, F.F. (1999). Response of coyote to avian distress calls: testing the startle-predator and predator-attraction hypotheses. *Behaviour*, *136*, 935-949.

Yorkston, E. & Menon, G. A. (2004). Sound idea: Phonetic effects of brand names on consumer judgments. *Journal of Consumer Research*, *31*, 43-51.