Sound Effects:

_Investigating the Cross-Modal Mappings Between Sound and Meaning_

B.A. Essay

Sara Sigríðar Ingvarsdóttir

Kt.: 130792-2439

Supervisor: Joe Jalbert

May, 2017
ぎりぎり [girigiri]

Japanese ideophone (Jisho, n.d.).

1. at the last moment; just barely (adj-na, adv)

Example: ぎりぎり間に合った

[girigiri ma-ni at-ta]

just-barely.ADV time-in.DAT arrive.PST

“I made it just barely in time”
Abstract

Language is generally assumed to be a strictly arbitrary system with no linkage between linguistic form and meaning other than convention. However, a growing body of research into sound symbolism has challenged the conventional view and demonstrated that people systematically map certain sounds with certain meanings. A well-known and widely attested example is the *bouba/kiki* effect. First observed by Köhler (1947), it describes how subjects of varying age and nationality consistently choose the pseudoword *bouba* for a rounded shape and *kiki* for a spiky shape. This paper provides an overview of the various types of sound-symbolic mappings that can occur in language and reviews pertinent experimental research on vowel-size correspondences and vowel-shape correspondences. Building on the theory that the articulatory feature of roundedness is the basis of vowel-roundness associations, this paper proposes a study to determine if there is a difference in sound-symbolic effects between individual rounded vowels. The study expands on the *bouba/kiki* paradigm and follows the experimental design of Ković, Plunkett and Westermann (2010), using a complicated label-object categorisation task to test the hypothesis that the pseudoword *búba* that has a tightly rounded vowel sound will produce a stronger roundness-association than the pseudoword *boba*, that has a less tightly rounded vowel sound. The prospective study is intended as a small step towards a growing understanding of the relationship between arbitrariness and sound symbolism in language.

*Keywords:* cross-modal correspondence, iconicity, linguistics, sound symbolism
# Table of contents

0 Abstract 0  
1 Introduction 1  
2 Background Information 2  
  2.1 Historical Context 2  
  2.2 Types of Sound Symbolism 4  
  2.3 Cross-Linguistic Variation 8  
3 Moving Beyond Debate 9  
4 Experimental Research 12  
  4.1 Sound-Size Correspondences 12  
  4.2 Sound-Shape Correspondences 14  
5 Possible Mechanisms 18  
  5.1 Acoustic Properties 19  
  5.2 Articulatory Properties 20  
6 Study Proposal 22  
  6.1 Vowel-Shape Associations 24  
  6.2 Hypothesis 25  
  6.3 Stimulus 27  
  6.4 Procedure 31  
7 Conclusion 33  
8 References 35
1 Introduction

The principle of *arbitrariness* in language describes the arbitrary connection between the sound of an utterance and its meaning (Perniss, Thompson, & Vigliocco, 2010), whereas the theory of *sound symbolism* implies a non-arbitrary relationship between sound and meaning (Hinton, Nichols, & Ohala, 1994). Sound symbolism is a controversial subject in linguistics because it conflicts with the assumption that language is strictly arbitrary. While not discounting arbitrariness, sound symbolism assumes that language is not categorically arbitrary and that words sometimes have a direct or indirect relationship with the meaning they express.

Because of its controversial status, most research on sound symbolism has focused on proving that sound-symbolic effects *exist*, but fewer studies have examined how these effects *work* (Lockwood & Dingemanse, 2015). There is bountiful behavioral evidence showing how people systematically map specific sounds with specific sensory meaning, such as *maluma* for a round shape and *takete* for a spiky shape (Köhler, 1947). However, the evidence is still inconclusive as to what drives the effects. To investigate the underpinnings of sound-symbolic effects, I will examine experimental research on cross-modal correspondences between the sound and meaning. I will review studies exploring sound-size mappings and sound-shape mappings. Furthermore, I will present the findings of Ohtake & Haryu (2013) and Oda (2000) that suggest sound-symbolic effects are connected to both acoustic and articulatory properties.

Lastly, I will propose a study to examine the sound-shape associations of rounded vowels with rounded shapes. I hypothesize that the articulatory feature of lip rounding is essential for associations with roundness in shape, and propose an experiment to test the this. To determine whether vowel-roundness associations differ between individual rounded vowels, the proposed experiment compares the strength of association between the two rounded vowels [u] and [ɔ] when separately contrasted with the unrounded vowel [i].
2 Background Information

In this section I will define key terms and provide background information on the development of arbitrariness and sound symbolism as theories of language. Beginning with an account of the pivotal ideas that have informed contemporary perspectives, I will note the first documented advocate of arbitrariness, Hermogenes from Plato's *Cratylus* (Lockwood, 2017; Sedley, 2003), and the most influential advocate of arbitrariness, Saussure (1916/1966). This historical context helps to explain how arbitrariness became the accepted theory and why the existence of sound symbolism has not been widely acknowledged until relatively recently (Lockwood & Dingemanse, 2015).

Having outlined the evidence used to preclude the recognition of sound symbolism, I will then describe sound-symbolic phenomena in more detail and consider the role of a Eurocentric perspective. First, I will review the main types of form-meaning mappings that can occur in language. Unpacking the typology is important to provide an entry point into the field and demonstrate the various types of relationships that can exist between sound and meaning. Second, I will examine cross-linguistic variation in the proportion of sound-symbolic words, arguing that a Eurocentric linguistic perspective has further hindered the acknowledgement of sound symbolism (Perniss et al., 2010).

2.1 Historical Context

The discussion of sound symbolism and arbitrariness as competing theories of language dates back centuries. Plato’s dialogue *Cratylus* contains a conversation in which two philosophers, Hermogenes and Cratylus, argue about the meaning of names and words (Sedley, 2003). They discuss whether words are “conventional” or “natural” in origin, i.e. whether they come to exist by convention and are thus arbitrary or whether they have some natural association with the entity they name. Hermogenes considers the relation completely arbitrary and argues that names are of conventional origin. Cratylus
disagrees. Socrates later intervenes in the discourse and provides a more integrated perspective, asserting that although many words have no apparent connection with their meaning, there exist “good” words where sound and meaning correlate. He concludes the dispute by remarking that while convention must be a factor, words are nevertheless always some kind of vocal imitations of the objects they refer to (Sedley, 2003).

Arbitraryness is the most adhered to theory today, commonly traced back to Ferdinand de Saussure’s (1916/1966) *Course in General Linguistics*. In it, Saussure stated that “The bond between the signifier and the signified is arbitrary” (p. 67) opposing the idea that words contain any intrinsic self-contained meaning. In his view, the existence of different languages suffices to disprove any natural connection, if sound-symbolic pairings did exist, they should hold cross-linguistically and not vary in form. Saussure's argument echoes that of John Locke (1690), who put it forth centuries earlier in his *Essay Concerning Human Understanding*. Locke argued that if sound and meaning had any natural connection, humans would all speak the same universal language. Since many different languages exist and they all have different words for the same ideas, he rejected the proposal of a motivated relationship entirely. Saussure advanced slightly further in his reasoning and considered the case of onomatopoeia, words that directly imitate their referent (e.g. *tick-tock, glug-glug*). He argued that onomatopoeic words “are only approximate and more or less conventional imitations of certain sounds” that evolve and become arbitrary linguistic signs once they have become established in the linguistic system (p. 69). His support for arbitrariness proved highly influential in linguistic theory and represents the conventional view of language, i.e. that words and sounds are arbitrary and no significant link exists between sound and meaning. The Saussurean view of considering language a strictly arbitrary system has dominated theories of language for over a century and linguists have only relatively recently begun to acknowledge the existence of sound symbolism (Lockwood & Dingemanse, 2015).
2.2 Types of Sound Symbolism

Form and meaning in language can correspond in a number of different ways and opinions vary on the optimal classification system as well as the scope of each category within it. The term *iconicity* is used as an umbrella term to cover the full range of linguistic signs that exhibit some sort of motivated correspondence between form and meaning (Perniss et al., 2010). Iconicity is widely observed in signed languages (Perniss & Vigliocco, 2014) and even animal communication (Hockett, 1959), but the term sound symbolism refers to iconicity in spoken language specifically (Lockwood & Dingemanse, 2015).

Discrete types of sound symbolism can be differentiated based on the degree of direct association between the sound of an utterance and its meaning. Hinton et al. (1994) suggest that the array of sound-symbolic phenomena is best described in terms of a continuum of form-meaning relationship. The continuum ranges from a completely arbitrary relationship between form and meaning, to a completely linked relationship. Each item's position on the continuum is determined by the extent of direct sound-meaning relation. Some sound-meaning associations have direct linkage, as in the case of onomatopoeic words that represent sounds in nature. Others have less direct but prominent relations, as when related forms are associated with related meanings, e.g. when a contrast between the vowels [i] and [a] produces a corresponding difference in judgements of magnitude (Lockwood & Dingemanse, 2015). And yet others sound-meaning associations have only abstract relations. Building on these distinctions, Hinton et al. (1994) propose a division of the overall concept of sound symbolism into four separate categories: corporeal, imitative, synesthetic and conventional sound symbolism. This classification is broadly equivalent to Von Humboldt's (1836/1999) description of sound systems as well as the more modern works of Cuskley and Kirby (2013) and Monaghan, Shillcock and Christiansen (2014). I will now present the main types of sound symbolism based on the categories defined by Hinton et al. (1994). Starting with the most direct sound-meaning associations, I will describe corporeal and imitative sound symbolism. I will then go on to discuss the more arbitrary conventional
sound symbolism. Finally, I will examine synesthetic sound symbolism, the type that is most relevant to this paper.

_Corporeal sound symbolism_ refers to voluntary and involuntary expressions of a speaker’s emotional or physical state through specific sounds or intonation patterns (Hinton et al., 1994). Because it represents a completely linked relationship between sound and meaning, it is positioned on the far end of the continuum. Involuntary symptomatic sounds like cries of pain, coughing and hiccuping belong to this category, along with interjections, expressive intonation and expressive voice quality. Vocative expressions, the use of vocalization to gain attention, fit into this category as well, e.g. purposefully clearing one’s throat to get attention. Corporeal sound-symbolic expressions generally occur in informal or unconventional speech and are scarcely seen in written form. Comic books are a rare exception because they regularly contain corporeal utterances, e.g. exclamations like _Achoo!_ or _Aaugh!_ and the shape and size of the letters is often manipulated to indicate intonation, loudness, and duration. However, corporeal sound symbolism differs from other types of sound symbolism because “the sound is not a true symbol, but rather a sign or symptom” (Hinton et al., 1994, p. 2). In other words, utterances of this type are not actual words or parts of words and do not carry the same kind of meaning as words do.

_Imitative sound symbolism_, or onomatopoeia, refers to phonetic sequences that directly imitate their referents through sound (Cruse, 1986). It also represents a very direct association with meaning but differs from corporeal expressions because it is verbal (Hinton et al., 1994). Typical examples of onomatopoeic vocabulary are words that denote environmental sounds, e.g. _bump, buzz_, and _whip_, and words that denote the sounds animals produce, e.g. _oink, meow_, and _moo_ (Perniss et al., 2010). It is semiotically the least complicated sound-meaning association and its existence has been recognized by even the most devoted supporters of the fundamental arbitrariness of the sign, such as Saussure and Newmeyer (Dingemanse, 2011). Similar to corporeal sound symbolism, onomatopoeia often represents sound patterns outside of conventional speech and is rarely encountered in writing. However, it is not directly rooted in
emotional experience like corporeal sound symbolism and has increased referential value in speech. Onomatopoeic words are much more frequently conventionalized into the lexicon compared with corporeal sound-symbolic utterances (Hinton et al., 1994). At the lexical level, onomatopoeia is the most conspicuous and the most acknowledged type of sound symbolism.

The type of sound symbolism called conventional sound symbolism refers to patterns of associations of certain phonemes and clusters with meanings. Hinton et al. (1994) note that this type resides closer to the arbitrary end of the continuum because the relevant sound-symbolic phonetic segments vary cross-linguistically. These meaningful segments, sometimes called phonesthemes (Bergen, 2004), exhibit relative iconicity because they depict related meanings. An example of a phonestheme in English is the onset “gl-” in words like glimmer, glisten, glitter, gleam, glow, glint, etc. This cluster does not frequently appear outside of words that mean something related to ‘light’ or ‘vision.’ The onset “sn-” is another example, occurring in many words relating to ‘mouth’ or ‘nose,’ e.g. snore, snack, snout, snarl, snort, sniff, sneeze, etc. According to Bergen (2004), phonesthemes seem to most frequently appear in content words rather than function words or more basic words. They are not exclusively onsets, e.g. the rime “-ash” in words like smash, bash, crash, and mash. Some words can even be interpreted as consisting of more than one phonestheme, e.g. the word sneer as a composition of both the previously mentioned onset “sn-” as well as the rime “-eer,” a phonestheme related to ‘contempt’ (e.g. leer and jeer). Bergen further reports that phonesthemes are prevalent in human language and have been found in the lexicon of all languages that have been systematically studied. Bolinger’s (1950) work on phonesthemes lends support to this claim, he identified and analyzed phonesthemic structures in English morphology and demonstrated that phonesthemes exist in far greater numbers than was previously recognized.

Synesthetic sound symbolism is the variety of sound symbolism that involves cross-modal sensory relations. Words belonging to this category consistently capture properties of their referents across modalities (Lockwood & Dingemanse, 2015).
Synesthetic sound symbolism sits close to the center of the continuum and Hinton et al. (1994) describe it as “the acoustic symbolization of non-acoustic phenomena” (p. 4). In other words, sounds that exhibit synesthetic sound symbolism carry with them a sense in perception, such as size, shape, brightness, manner of movement etc. Lockwood (2017) provides a thorough overview of the main findings. Put simply, low back vowels (e.g. [a], [o]) and voiced consonants (e.g. [b], [g]) are associated with round, large, slow, dark and heavy things. And small vowels (e.g. [i], [e]) and voiceless consonants (e.g. [p], [k]) are associated with spiky, small, quick, bright and light things. Additionally, low vowels and back vowels are associated with big things, whereas high vowels and front vowels are associated with small things.

Ideophones (also known as mimetics or expressives) are sound-symbolic words that exemplify synesthetic sound symbolism. Dingemanse (2012) comprehensively defines ideophones as “marked words that depict sensory imagery” (p. 655). That is, ideophones are marked in the sense that they are unmistakably distinct from other words in a given language. They are words as opposed to merely sounds. They depict meaning rather than describe it, and the meaning they depict is sensory, i.e. derived from sensory perceptions. They commonly depict sensations such as smell, vision, touch, sound and body movements. For example, the ideophone sarasara refers to something that is smooth to the touch, whereas zarazara refers to a rough surface (Watanabe & Sakamoto, 2012). Ideophones have similarities with onomatopoeia but the two are not interchangeable. Onomatopoeia refers to the direct mapping of a sound pattern to a word, whereas ideophones are not limited to just auditory events (Cuskley & Kirby, 2013). They have also been grouped with phonesthemes, but ideophones are recognizably different in that native speakers are highly conscious of their sensory motivation. They seem to stand out from other words in a lexicon and have been described as being striking and phonologically peculiar. Dingemanse (2012) states that they are often special word forms, have foregrounded prosody, expressive morphology and relative syntactic independence.
2.3 Cross-Linguistic Variation

Indo-European languages generally contain fewer sound-symbolic words and Perniss et al. (2010) suggest that a Eurocentric linguistic perspective is the reason why sound symbolism has not received much attention in mainstream scientific research. The inventory of iconic forms in language has commonly been described as insignificant, and that might be accurate for most Indo-European languages. Indeed, the lexicon of English shows little evidence for anything but conventionally determined arbitrary connections between words and their meanings. However, Perniss et al. emphasize that other language families and signed languages have widespread iconic mappings. Ideophones are prolific in the languages of Africa, Asia and the Americas where they have been extensively researched (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015).

Although Newmeyer (1992) and Saussure (1916/1966) granted the existence of sound symbolism, they both dismissed it as only a negligible part of the vocabulary in any given language. In Newmeyer's words, “The number of pictorial, imitative, or onomatopoetic non-derived words in any language is vanishingly small.” (p. 758). Perniss et al. (2010) reject his claim, pointing out how well documented iconicity is throughout different levels and modalities of language and its growing recognition as a general attribute of language. The presence of motivated sound-meaning relations becomes especially apparent when languages outside of the Indo-European family are examined (Imai, Kita, Nagumo, & Okada, 2008).

In some languages ideophones even form recognizable word classes with distinct phonological, morphological, syntactic and semantic properties. For example, Diffloth (1976) analysed the Semai language (an Aslian language spoken in Western Malaysia) and reported that ideophones in Semai constitute a third basic word class that is equal to nouns and verbs in size. Ideophones also constitute a large word class in the Niger-Congo languages Bantu (spoken in Central Africa, Southeast Africa and Southern Africa) and Gbeya (spoken in Central Africa) (Samarin, 1971).
3 Moving Beyond Debate

Arbitrariness and iconicity have long been discussed as if they were mutually exclusive theories, but this an overly simplified interpretation. To provide a more constructive perspective, this section will explore the advantages that each system brings to communication and word learning. Arbitrariness is recognized as a general characteristic of language and I will argue that sound symbolism should be recognized in the same manner. Recent works of research have supported the recognition of sound symbolism as a general property of language (Monaghan & Christiansen, 2006; Monaghan, Shillcock, Christiansen, & Kirby; 2014) and based on this evidence, I reason that they are best defined as two complementary systems of language.

Arbitrariness is a fundamental assumption in theories of language processing and many psychological and neuroscientific models of language have been organized around it. It is recognized as an inherent element in language, considered by Hockett (1960) as one of the 13 key components that characterize human language and distinguish it from animal communication. In Hockett’s view, “The design feature of ‘arbitrariness’ has the disadvantage of being arbitrary, but the great advantage that there is no limit to what can be communicated about.” (p. 6). That is, while he concedes that the unrestrained nature of arbitrariness is disadvantageous, he argues that it allows a larger lexicon and greater range of communication.

Likewise, Gasser (2004) argues that arbitrariness is critical for language to encompass so many concepts without ambiguity or overlap in meaning. Gasser observed an interaction between learning strategy and vocabulary size that could explain why human language is inclined towards arbitrariness. He simulated word learning in a computational model and demonstrated the long-term advantages of arbitrariness for acquiring a large-scale vocabulary. Gasser designed two sets of form-meaning pairings for the model to learn, one set had arbitrarily associated pairings and the other had iconically associated pairings. He trained the model on each set
separately to compare how effectively it could acquire the associations between the forms and meanings. Despite an early advantage, iconic form-meaning associations significantly restricted the size of vocabulary that the learning algorithm acquired over time. The arbitrary form-meaning associations granted the model more flexibility to form mappings and resulted in superior performance when many patterns had to be acquired.

Considering this, Gasser (2004) hypothesizes that the proportion of iconic vocabulary in a language varies with the overall number of words in its lexicon. His work is important to consider since much of the empirical evidence for sound symbolism comes from behavioral studies that center on word learning. The results of Gasser's simulation show the superiority of a purely arbitrary word-learning strategy over a purely systematic one. He argues that this superiority is the reason why form-meaning relations in language are mostly arbitrary, the sheer number of concepts human language can represent necessarily make it rely on arbitrary relations. But even so, his findings do not imply that iconicity is non-existent in language. Given a mixed set of form-meaning pairings where some are arbitrarily related and some are iconically related, Gasser notes that a more complicated model would successfully make use of both strategies. which could explain why iconic mappings persist in language.

Despite the clear benefits arbitrariness brings to word learning and usage, it does not exclude sound symbolism. In fact, recent research on language acquisition provides evidence for a complementary analysis of arbitrariness and sound symbolism. Monaghan and Christiansen (2006) found evidence for both arbitrariness and iconicity by analyzing English and French corpora. They simulated word learning in a neural network model and found advantages for arbitrary mappings, replicating Gasser's (2004) findings. However, they also observed distinct benefits for iconic mappings in learning what category a word belongs to. Based on their findings that both arbitrariness and sound symbolism bring advantages to word-learning, they claim that the cohabitation of the two systems is a design feature in language that facilitates the language acquisition process.
This view is supported in a later study by Monaghan, Shillcock, Christiansen and Kirby (2014), who demonstrated pronounced iconicity in the vocabulary of children in a large-scale corpus analysis of English. Monaghan et al. analyzed the sound-meaning relations in 300 one-syllable words that children acquire before the age of 4 years old, and found that words acquired early were more iconic than words acquired late. They argue that sound symbolism is proportionally high in early vocabulary because it supports language learning, with arbitrariness incorporated for communicative efficiency in later language. Furthermore, through their analysis of English they found that sound-meaning relations in English vocabulary are more systematic than can be expected by chance. Based on their results, they argue that sound symbolism is not a negligible part of the vocabulary, but rather a general property of language.

A complementary view of arbitrariness and iconicity is supported by Perniss et al. (2010) who propose a model of language in which both systems are essential. They hypothesize that two fundamental constraints influenced the evolution of language, one that favors arbitrariness and the other iconicity. In their view, the cross-linguistic variability in the proportion and form of iconicity shows how every language balances these constraints differently. The first constraint is “the need to ensure an effective linguistic signal” (p. 11), lending support for arbitrariness as a necessary property of language. As previously noted, arbitrariness is vital to the development of a large lexicon because it allows for greater distinction between each lexical entry (Gasser, 2004). For this reason, Perniss et al. state that arbitrariness contributes pragmatically to the effectiveness of communication in that communicative success would be greatly diminished if all words in a particular semantic domain sounded very similar.

The second proposed constraint is “the need to link linguistic form to human experience” (p. 11), where iconicity becomes important. Perniss et al. (2010) argue that this requirement is essential to the embodiment theory of language. Embodied views of language postulate a connection between language processing and actions. According to
embodiment theory, language comprehension, for example, requires a mental re-enactment of the specific embodied experience, thus engaging the motor system in the brain during processing. Based on embodied views, Perniss et al. claim that the same systems used in perception and action are triggered by language use (i.e. comprehension, production, and acquisition). They further remark that iconicity was essential to the evolution of primitive communication into complex linguistic systems. According to them, both vocal and manual gestural forms were evolutionary precursors to the development of human language. The synchronous evolution of both types of gestures supposedly enabled a symbolic system to emerge based on the relations between the physical world and representational forms. Perniss et al. contend that the integration of linguistic and imagistic systems suggests a linkage from the beginning. Gestures have been argued to be manifestations of the mental re-enactments suggested by embodiment theory.

4 Experimental Research

The most widely studied sound-symbolic associations are probably sound-size correspondences and sound-shape correspondences. These associations are also the most relevant to this paper as my study proposal examines vowel-shape correspondences. In the following two sections I will examine studies on vowel-size associations by Sapir (1929) and Lockwood, Hagoort and Dingemanse (2016). Next, I will discuss Köhler's (1947) vowel-shape study and several replications and adaptations of his experiment.

4.1 Sound-Size Correspondences

The correlations between vowel sounds and object size have been extensively studied in sound-symbolism research. The experiments conducted by Sapir (1929) are well known demonstrations of this specific mapping. Sapir presented his participants with two
tables, one small and one big, and asked them to judge which label, *mal* or *mil*, referred to the larger table. He structured his pseudowords to only contrast the vowel sound, the back vowel sound in *mal* and the front vowel sound in *mil*. He found that 80% of participants chose *mal* as the large-sounding label, and this pattern held over many different pairs of words exploiting this front/back distinction.

In a study using synthesized size/sound sound-symbolic pseudowords, Lockwood et al. (2016) included a neutral condition to explore whether vowel-size correspondences are a graded effect based on the front/back distinction. They expanded on the traditional binary forced-choice paradigm and devised a vowel-size experiment with big-sounding pseudowords (e.g. *badobado*), small-sounding pseudowords (e.g. *kitikiti*), and neutral-sounding pseudowords (e.g. *kedekede* for big and *depedepe* for small). The experiment consisted of three conditions: a match condition where participants learned the congruent meaning of sound-symbolic pseudowords (e.g. *badobado* as big), a mismatch condition where participants learned the incongruent meaning of sound-symbolic pseudowords (e.g. *tipitipi* as big), and a neutral condition where participants learned neutral words that neither matched nor mismatched with their meaning (e.g. *kedekede* as big). Lockwood et al. included the neutral condition to have a comparison for the sound-symbolic pseudowords. This allowed them to ascertain whether the effect for vowel-size associations was a mapping boost because of the cross-modal congruency in the match condition, a mismatch hindrance from the cross-modal clash in the mismatch condition, or a combination of both.

Lockwood et al. (2016) reviewed the potential results of their study in terms of a match boost effect, a graded effect or a mismatch effect. They reasoned that if participants learned matching pseudowords (e.g. *badobado* as big) better than neutral pseudowords (e.g. *kedekede* as big), and neutral pseudowords better than mismatching pseudowords (e.g. *badobado* as small), that would indicate that the effect is a graded one. However, if they learned matching pseudowords better than neutral pseudowords, but there was no difference between neutral pseudowords and mismatching pseudowords, that would indicate a match boost effect. Lastly, if the participants
learned the mismatching pseudowords worse than matching or neutral pseudowords, but there was no difference between matching and neutral pseudowords, that would imply a mismatch hindrance effect. This is where the neutral condition is useful because it allows them to discriminate between an actual cross-modal clash and a simple lack of cross-modal correspondence.

Native Dutch speaking participants went through three learning rounds in which they learned to classify the 36 pseudowords as either big or small. They were not split into groups and all participants learned the same set of words and meanings. They sat in front of a computer screen and saw either the Dutch word *groot* (big) or the word *klein* (small) appear on the screen, followed by a blank screen while a pseudoword was presented auditorily. A final screen showed the Dutch meaning with the pseudoword in written form (e.g. *groot* + *badobado*). Participants did not give any feedback in the learning round but pressed a button to move on to the next word.

In a subsequent test phase, the Dutch word *groot* or *klein* again appeared on the screen, followed by a question mark while a pseudoword was presented auditorily. Participants were then asked to identify whether the pseudoword matched the Dutch word by pressing a button. They identified the matching pseudowords correctly 75.56% of the time, the neutral pseudowords 66.11% of the time and mismatched pseudowords 62.50% of the time. Lockwood et al. (2016) infer from these results that the effect is a matching boost. Participants learned sound-symbolically congruent pseudowords (e.g. when *badobado* meant big and *kitikiti* meant small) better than the neutral pseudowords (e.g. when *kedekede* meant big and *depedepe* meant small), but there was not a distinguishable difference in how well they learned the matching/mismatching neutral pseudowords.

### 4.2 Sound-Shape Correspondences

The widely used *bouba/kiki* experimental paradigm comes from a well known study by Köhler (1947) on the cross modal associations between sound and shape. Köhler found
that Spanish-speaking participants consistently linked "spiky" words like takete with spiky shapes, and "round" words like maluma with round shapes. The effect came to be known as the bouba/kiki effect because of Ramachandran and Hubbard's (2001) replication with the labels bouba and kiki. They found that 95% of English speakers linked bouba with a round shape and kiki with a pointy shape. The bouba/kiki effect was even observed in a study on the Himba, an indigenous tribe in Namibia who have no written language or exposure to Western culture (Bremner, Caparos, Davidoff, de Fockert, Linnell, & Spence, 2013). Ramachandran and Hubbard theorize that the bouba/kiki effect is the result of a cross-modal association between visual features (round or pointy) and articulatory aspects of sounds. They argue that a resemblance between the shape of the lips when you produce a word corresponds somehow with the object the word refers to, the word will be more strongly linked to its meaning.

Maurer, Pathman and Mondloch (2006) replicated Ramachandran and Hubbard's (2001) results in a study with both children and adult participants. Participants were presented with two images, and a choice of pseudowords such as bouba/kiki, takiti/mabuma. They found that both children and adults systematically mapped round shapes with rounded vowels and spiky shapes with unrounded vowels. Moreover, in a follow-up experiment, Spector and Maurer (2013) used the same paradigm with better controlled stimuli. They point out that Ramachandran and Hubbard's labels varied in both consonant context (e.g. bouba vs. kiki) and syllable variability (e.g. kiki vs. maluma) and do not provide an adequate contrast to investigate the independent influence of vowels. To address this, they created new pairs of words that were only differentiated by the vowel. The words contrasted the rounded back vowel [o] and the unrounded front vowel [i] in the context of a reduplicated stop consonant, e.g. gigi/gogo, bibi/bobo, etc. As in the first experiment, the children associated the rounded vowel [o] with rounded shapes and the unrounded vowel [i] with spiky shapes significantly more often than could be predicted by chance. The effect was present across all consonant contexts, but was slightly stronger for the rounded vowel-shape mapping.
These findings demonstrate that children, like adults, show sound symbolism biases (Maurer et al., 2006; Spector & Maurer, 2013). Based on these findings, Spector and Maurer (2013) speculate that sound-shape mappings are a natural bias that may facilitate language acquisition. The findings are also important in ruling out orthography because the children participating in both studies were 2-3 years old and could not yet read. Although 'O' is a curved letter and 'I' is an angular letter, any orthographic interference can effectively be ruled out because the children were not familiar with the orthographic representations of the vowels.

Ković, Plunkett and Westermann (2010) examined the round/spiky association using a more complicated matching task. Their study consists of two behavioral experiments in which participants performed language tasks. In the first experiment the reaction times of participants were recorded, in the second experiment reaction times and ERPs were recorded. The first experiment was a novel word learning experiment with 37 participants who were all native speakers of English. For visual stimuli, Ković et al. used schematic drawings of objects that were composed of four binary valued features, relevant for the round/spiky distinction. For auditory stimuli, they recorded the feedback phrases 'It's a mot!' and 'It's a riff!' These labels were designed to be round-sounding (mot) and spiky-sounding (riff) based on Maurer et al. (2006). Beginning with a training phase, participants learned to assign a label to an object based on either a congruent (sound-symbolic) condition or an incongruent (non-sound-symbolic) condition. They saw an object on the screen and pressed a button to indicate whether the object was a mot or a riff. If they got it right they heard a chime sound followed by the correct label in a feedback phrase ('It's a mot/riff'), if they got it wrong they heard a buzzer sound and a feedback phrase. In the following test phase, participants were presented with a label/object pairing and asked to press either a 'match' or a 'mismatch' button as fast as they could identify a match/mismatch.

The second experiment in their study was very similar to the first. The same visual stimuli were used, there was a training phase and a test phase. 31 native English
speakers participated, split into two groups for congruent/incongruent conditions. In the training phase, one group learned sound-symbolically congruent labels for round and spiky objects (i.e. *dom* for a round object and *shick* for a spiky object), and the other group learned incongruent labels (i.e. *shick* for a round object and *dom* for a spiky object). In a test phase following the training phase, participants heard the name for a label and saw an object on the screen. As in the first experiment, their task was to determine whether the object matched with the label or not by pressing a button.

Ković et al. (2010) found effects for sound-symbolic label-object associations, but they were much stronger for roundedness than pointedness. The behavioral data showed that participants in the sound-symbolically congruent condition were significantly faster to identify an object-label match and reject a mismatch, compared with the group in the incongruent condition. Rounded object/label pairings resulted in faster performance than pointy object/label pairings. Furthermore, objects with congruent sound-symbolic labels (like *dom* for a rounded object) elicited a stronger negative-peaking wave between 140-180 ms after presentation than objects with incongruent labels. This effect was seen at the O1 and O2 sites, i.e. the occipital region where visual processing occurs. Ković et al. speculate that this early negativity represents the brain's response to a mapping between auditory and visual features, based on previous research by Giard and Peronnet (1999) and others. Furthermore, they detected mismatch effects for roundedness in both congruent and incongruent condition. The ERP signatures were distinctly different when participants accepted the round-sounding label *dom* for a round object in the congruent condition, or rejected it for a round object in the incongruent condition. Ković et al. theorize from this that the mapping between a rounded object and a rounded label pairing is treated differently in processing. Both the behavioural and neuroimaging results show more robust effects for roundedness than pointedness. Ković et al. speculate that either their stimuli was not pointy enough or that pointy shapes are simply weaker sound-symbolically than round shapes.
Alternatively, Lockwood (2017) suggests that roundness is more salient because associations in the round/spiky domain are dominated by perceptions of roundness. He points out that we need to be aware of the limits in what the bouba/kiki experimental paradigm can tell us about sound symbolism. In the case of sound-shape correspondences in the vowel domain, the binary forced-choice task showed that people systematically correlate rounded vowels with round shapes and unrounded vowels with spiky shapes. Although this paradigm is useful for establishing the existence of effects, it cannot evaluate the role of each association separately because there are only two words being contrasted. This has resulted in the assumption that round/spiky associations are equally strong. Results of more complicated experiments such as the study by Ković et al. (2010) are increasingly showing that this assumption is wrong. A more accurate analysis, related by Lockwood, is that rounded sounds have a strong connection with can create an artefactual connection between unrounded sounds and spiky shapes when there are only two words being contrasted.

5 Possible Mechanisms

Both acoustic and articulatory properties have been hypothesized to be the mechanisms of sound symbolism and the ensuing discussion about which of the two is the main driver dates back to the behavioral research of Sapir (1929) and Newman (1933). Neither Sapir nor Newman argued strongly for either property, but rather acknowledged the involvement of both. The following sections will examine two pertinent research papers that looked at both properties in an experimental setting. First, I will present findings that support an acoustic mechanism from a study by Ohtake and Haryu (2013). Next, I will present evidence from Oda's (2000) studies, suggesting an underlying articulatory mechanism. The exact function of articulatory and acoustic properties is still unclear, but these findings raise intriguing questions regarding the nature and extent of their involvement.
5.1 Acoustic Properties

Acoustic properties of sounds have been argued to be the motivating factor behind sound-symbolic effects. Ohtake and Haryu (2013) explored the role of acoustic features in a study building on the widely researched front/back vowel distinction. They conducted two experiments with native speakers of Japanese to investigate the underlying processes of vowel-size correspondences. Their study makes a novel contribution to the field by considering articulatory and acoustic factors separately.

Ohtake and Haryu's (2013) first experiment focused only on acoustic properties and involved no articulation. In a classification task, participants were presented with an image of a disk on a screen and asked to judge its varying size, "big" or "small," with a press of the button. Out of a total of 60 trials, the sound [a] played during 20 trials and the sound [i] for another 20. The experiment was designed to examine whether reaction times to the varying size of the image were influenced by the acoustic sound playing in the background. Indeed, this was the case, participants were significantly slower to respond in the incongruent condition when [a] was presented with a small image or [i] with a big image. These results are consistent with those obtained by Sapir's (1929) mil/mal experiments, showing an association between [a] and big, and [i] and small.

Conversely, the second experiment focused only on articulatory properties with no accompanying auditory experience. Ohtake and Haryu (2013) designed this experiment to isolate articulation by involving the proprioception of the size of the oral cavity. In the same classification task as in the first experiment, participants judged the varying size of an image on a screen while keeping either an egg-shaped or a board-shaped object in their teeth. This method was used to simulate tongue position and mouth openness when pronouncing the two target vowels. The experimental trials comprised six blocks of 72 trials, divided into three phases. In the first phase participants held the egg-shaped object in their teeth, in the second phase they held the board-shape object, and in the third phase they held nothing. As in the first experiment, reaction times were recorded to compare between conditions. In this case there was not a significant difference between reaction times in the congruent and incongruent
conditions.

The results of Ohtake and Haryu's (2013) experiments suggest that the systematic size correspondences entailed in the vowel sounds [a] and [i] are attributable to acoustic rather than articulatory features. The first experiment demonstrated that the acoustic features of the vowels elicit perceptions of a smaller size for [i] and bigger size for [a]. The second experiment indicated that articulation on its own may not be sufficient to elicit a correspondence with perceptions of size. However, the proprioception method Ohtake and Haryu used to mimic articulation was limited to a static articulatory experience, rather than a dynamic one. Participants did not move their lips and tongue as they would during speech. The actual movement of the mouth could contribute to the vowel-size correspondences and the experiment does not account for that. Although Ohtake and Haryu cannot rule out the contribution of articulatory features, they conclude that acoustic features have a primary function in vowel-size correspondence.

5.2 Articulatory Properties

Articulatory properties of sounds have also been put forth as the mechanism behind sound-symbolic effects. In contrast to the study by Ohtake and Haryu (2013), Oda's (2000) research indicates that articulatory properties are a key factor in establishing form-meaning mappings. In Oda’s (2000) study, participants that pronounced Japanese ideophones themselves were significantly more accurate at correctly guessing the English definitions of ideophones compared with the group that did not.

Oda's (2000) experiment consisted of a questionnaire in two parts. The first part was composed of 15 questions. For each question participants were presented with a cartoon of a situation, a short English description of the situation and three ideophones. One ideophone actually described the situation, one was opposite to it in some way, and one was unrelated. The words were positioned below the cartoon in random order and participants had to choose which one they felt best fit the situation. For example, one
question included a cartoon of a jumping rabbit along with the description "Hopping or skipping agilely" and the three ideophones *pyonpyon* ("jumping"), *gosogoso* ("moving around restlessly") and *noronoro* ("sluggish, moving around slowly") (p. 319).

The second part of the experiment was composed of 5 questions. In each question, two cartoons/explanations such as the ones from the first part and two minimal-pair ideophones were presented to the participants. For example, one question included a cartoon of an electrocuted person holding a clothes iron along with the description "That feeling one gets from an electric shock or such," a cartoon of a woman sunbathing and grimacing from the heat on her skin with the description "tingling in pain on the skin or on the tongue," and the two ideophones *biribiri* ("electric shock") and *hirihiri* ("tingling pain") (p. 321). Oda (2000) chose the binary-choice format because he predicted that this would be a significantly harder task compared with the first part.

Participants were divided into two groups for both parts of the experiment. The 21 participants in the so-called "sound group" listened to a native Japanese speaker read the words out loud twice and then had to focus on the sound before making a decision on the questionnaire. The 49 participants in the "articulation group" likewise listened to a native speaker read the words, but were then prompted to pronounce the words themselves before answering the questions. They did not receive any guidance on how to pronounce the words, but were simply asked to pronounce them once after the speaker (Oda, 2000).

While both groups performed at above-chance accuracy, the articulation group outperformed the sound group in both parts of the study. Pronouncing the words seemed to give participants in the articulation group an advantage for choosing the correct ideophones. These findings suggest that articulation may be a significant factor in sound symbolism. Oda (2000) argues that all speakers produce relatively similar sounds and thus share similar articulatory experiences. In his view, these shared articulatory sensations enable non-native speakers to guess the meanings of sound-symbolic words. He further speculates that the cross-linguistic differences between articulatory
sensations (such as tongue placement for the [t] sound in Japanese and English) explain why there are language-specific variations in sound symbolism. One question that needs to be asked, however, is whether the experience of articulation is the only meaningful difference between the two participant groups. The act of articulating the words might have resulted in better performance because it required a deeper engagement with the task. Oda fails to address this potential confound.

Despite relatively sparse descriptions of methods, Oda (2000) seems to have observed robust sound-symbolic effects overall. Over 90% of participants from both groups correctly identified the ideophones buruburu ("shivering") and pyonpyon ("jumping") and Oda (2000) argues that these labels have especially transparent articulatory connections with meaning. He deconstructs the articulatory features of the ideophone pyonpyon into three parts. First, he connects the initial plosive sound 'p' where the lips are closed and tensed with the action of bending the legs to prepare to jump. Second, he connects the 'yo' sound in which the air is released and the tongue is thrust forward with the actual jump. Third, he connects the 'n' sound in which the tongue moves towards the back of the mouth and movement ceases with the landing. Oda theorizes that the sound group was able to correctly guess the meaning of these words because they represent easy cases with transparent connections to articulation. In such cases, he argues that speakers do not need to pronounce the words to realize their meaning, their past experience of producing sounds and body movement suffices to establish the connection. In other cases, the connection between meaning and articulation is more subtle and Oda (2000) asserts that the experience of pronouncing the word aids performance.

6 Study Proposal

Studies of sound symbolism have demonstrated that people of diverse linguistic backgrounds consistently associate certain sounds with certain shapes, such as the labels bouba/kiki and maluma/takete with rounded/sharp features (Köhler, 1947; Maurer,
However, it remains to be answered how these associations work, and what mechanism underlies them. To answer this question, many have theorized that sound-shape associations are grounded in some resemblance to articulatory experience, i.e. what it feels like to pronounce the sounds (Jones, Vinson, Closter, Zhu, Santiago, & Vigliocco, 2014; Ramachandran & Hubbard, 2001; Oda, 2000). Based on that theory, I propose the following study to seek further evidence for the articulation hypothesis and more accurately assess the connection between articulatory properties and sound-symbolic effects.

The study is intended to contribute to the growing area of sound-symbolism research by exploring roundness associations in the vowel domain. The primary objective is to determine whether vowel-roundness mappings differ between individual rounded vowels. Although extensive research has been carried out on sound-shape associations, no previous study that I know of has examined how the effects vary when accounting for individual rounded vowels. A secondary aim of the proposed experiment is to investigate if sound-symbolic preferences can be demonstrated with speakers of Icelandic. Although I fully expect to find significant effects based on findings from other Germanic languages, finding effects for Icelandic speakers would lend further support for the theory that sound symbolism is a universal feature of language.

In the following section, I will discuss what others have theorized about the relationship between rounded sounds and round shapes. Next I will present my hypothesis, that the articulatory feature of lip rounding is the driver of the association. I will then describe the experiment designed to test this hypothesis, discussing stimulus design and methodology based on the study by Ković et al. (2010). Afterwards I will discuss the result I predict based on my hypothesis, that a tighter rounded lip position (as in the vowel [u]) yields a stronger association to roundness than a more relaxed position (as in the vowel [ɔ]). I will also discuss the alternative results, that either the vowel [ɔ] has a stronger association, suggesting that vowel height is more important for roundness, or that there is no significant difference between the two vowels. Lastly, I will discuss implications for further research. The present study can tell us whether
there is a difference in effect between the two rounded vowels, further studies can build on the results and explore potential differences in more detail.

6.1 Vowel-Shape Associations

Sound-symbolic correspondences are not homogeneous and operate in various ways. For example, many of the sound-symbolic associations between vowels and sensory meanings seem to map onto a scale based on vowel height. This scale goes from the front high vowel [i] to the mid/low back vowels [a] and [o], and maps onto perceptions of small size/large size, fast speed/slow speed, lightness/heaviness, and brightness/darkness. Meanwhile, consonant voicing seems to be a binary distinction that maps onto the extreme end of the same scale; i.e. voiceless/voiced consonants map onto light/heavy, bright/dark, fast/slow, and small/large (Lockwood, 2017).

Whereas the front/back mappings are fairly well understood, considerably less is known about the mappings between sounds and shapes. Köhler’s (1947) bouba/kiki effect for round/spiky shapes has long been considered a scaled effect like the front/back distinction of consonant-voicing. However, recent findings have challenged this theory with more complex experimental paradigms, demonstrating significant effects for round shapes, but not for spiky shapes (Jones et al., 2014; Ković et al., 2010). These findings suggest that round/spiky associations do not map onto a graded spectrum, or at least it they are not driven equally by both ends of the spectrum. Rather, there is a strong association for roundness that has created an incidental association with spikiness in binary forced-choice experiments (Lockwood, 2017; Jones et al., 2014).

Several characteristics of sounds have been associated with sensory perceptions of object roundness. The sounds that have consistently been associated with round shapes are: labial consonants, voiced consonants, rounded vowels and low-back vowels (Lockwood, 2017). Following the hypothesis that the articulatory properties of sounds are the basis of sound-symbolic associations, that leaves me with several options to examine. However, since there is not a spectrum from round to spiky, I cannot compare
the far ends of the scale to form a hypothesis on which articulatory property is important.

In the vowel domain, roundedness seems to be an instrumental property. Jones et al. (2014) argue that sound-shape mappings are based on the articulatory rounding of the lips through bilabial plosives and rounded vowels. In their view, a rounded lip posture presents a visual and motoric connection to roundness. They argue that this is the reason why roundness dominates the round-spiky domain, because there is no similar connection we can point to for spikiness.

### 6.2 Hypothesis

Following the theme that rounded vowels carry with them a strong perception of roundness in shape, I propose the present study to investigate individual differences between two different rounded vowels. The purpose of this comparison is to examine whether the associations between vowels and round shapes are a graded effect similar to the front/back distinction of vowel-size associations, or a binary distinction. To accomplish this, I designed a set of vowel-contrasts to examine the effects of lip rounding on roundness associations. I contrast the rounded vowel [u] with the unrounded vowel [i], and the rounded vowel [ɔ] with the same unrounded vowel [ɨ] to determine if there is a meaningful difference between their sound-shape effects.

I hypothesize that roundness associations are grounded in the articulatory property of lip-rounding, and that the degree of lip rounding determines how strongly the sound is associated with round shapes. Based on this hypothesis, I predict that the more tightly rounded [u] vowel sound will produce a stronger roundness-association and faster response times compared with the less tightly rounded vowel sound [ɔ]. Consequently, I predict that participants in the sound-symbolically congruent conditions will be faster to accept a correct label-object match and reject a mismatch, and I expect this effect to be stronger for [u] than [ɔ].
Alternatively, if I find that the vowel sound [ɔ] is more strongly associated with roundness than [u], I can speculate that differences in height between the two vowels are contributing to the effect. Vowel height is instrumental for many sound-symbolic associations and differences in height between rounded vowels could determine the strength of the roundness association. Based on this theory, I should see a stronger association with the mid vowel [ɔ] compared with the high vowel [u] because low-back vowels have been connected with roundness (Lockwood, 2017).

A more open and relaxed rounded lip posture might even be a stronger cue for roundness than a tight posture. Roundedness is rarely defined as anything but a binary feature of vowels and it is unclear what constitutes a more rounded sound or a less rounded sound articulatorily. Based on the same claim that lip rounding is the main mechanism behind roundness associations, the graded effect might be reversed; stronger for a slack lip posture and weaker for a tight lip posture. Based on this, the vowel sound [ɔ] is potentially more "round-sounding" than [u] because it has a more open and slack articulation.

However, different degrees of lip rounding and protrusion may not even make a significant difference to the effect. Roundedness may simply be a feature of rounded vowels that gets people to the semantic domain of roundness, but differences in lip posture may not make a difference for the effect. Rather than a graded distribution from a rounded lip posture to a spread lip posture, vowel-roundness associations could be a binary distinction in which rounded vowels are systematically linked with round shapes and unrounded vowels are not. In that case, I would not observe a significant difference between the vowels.

Although this experiment cannot answer whether it is vowel height or lip rounding that drives the effect, it provides a starting point for further research. Because I only compare two vowel sounds and their associations with roundness, the results will only tell us whether there is a difference between their effects, and not what potential
articulatory properties are responsible. The two sounds differ in lip posture, tongue position and height and the results will be inconclusive as to which of these differences is important. However, follow-up studies could address the confounding variables and examine the roundness association in more detail. For one, the study can be repeated with different rounded and unrounded vowels to examine how the vowel space is mapped to shape. It would also be interesting to assess the involvement of vowel height by comparing the low-mid vowel [œ] with the high vowel [u]. A further study could also examine orthographic effects by contrasting the more curvy letters O/Ö ([œ] and [œ]) with the less curvy letters U/Ú ([y] and [u]). Furthermore, when we have more information about what is important for roundness associations, we can potentially find a neutral condition, i.e. a vowel sound that is "halfway" associated with a round shape. A neutral condition would enable us to replicate for the vowel-shape domain the experiment that Lockwood et al. (2016) conducted with sound-size correspondences, and investigate how roundness associations affect learning.

Taking into account Lockwood's (2017) criticism of binary forced-choice tasks, I emphasize that this study is only intended to determine if there is a significant difference between the strength of roundness associations elicited by the two vowels. The unrounded vowel [i] is not meant as a spiky condition and I will not make inferences based on the potential mapping between [i] and spiky visual stimuli.

6.3 Stimulus

The stimuli of the study consist of the three pseudoword labels *boba, búba* and *boba* presented both auditorily and visually on a screen, and schematic animal-like figures presented visually.

The visual stimuli are borrowed from Ković et al. (2010). They used schematic drawings of animal-like figures as their visual stimuli (see Figure 1). Each figure is composed out of four binary-valued features; a rounded/pointy head, crescent/triangle wing, rounded/bushy tail and three/five legs. They designed the drawings based on
similar figures by Rehder and Hoffman (2005) and the popular 5–4 categorization model by Medin and Schaffer's (1978), where features varied in relevance based on where people are most likely to look. The head and tail of the figure were always either round or non-round in shape, whereas the wing and legs were not as contrastive. This design assumes that participants will pay most attention to the head, resulting in 77% accuracy if they answer only based on the head shape. If they only paid attention to the legs or wing, however, they would perform at chance level accuracy. Ković et al. reason that this categorization model elicits the same processes that are involved in natural language interpretation.

Figure 1: Visual stimuli used by Ković et al. (2010).

The inventory of Icelandic vowels contains four rounded vowels out of a total of eight monophthongs. There rounded vowels are the two front vowels: [ʏ] (as in 'um’) and [œ] (as in 'glid’), and the two back vowels: [u] (as in 'ur’) and [ɔ] (as in 'oft’). Their articulatory properties are commonly described in three main ways. First, they are classified in terms of backness, ranging from front to back depending on whether the tongue is positioned towards the back or the front of the mouth. Second, they are described in terms of height, ranging from low to high depending on the tongue's vertical position in the mouth. Third, they are either rounded or unrounded, depending on whether the lips are rounded (Rögnvaldsson, 2013). Additionally, rounded vowels have either compressed or protruded lip posture (Roger, 2014). Although roundedness is usually described as a binary feature of sounds, the lip protrusion/compression distinction is a little more descriptive of the types of rounding that are possible. Table 1
summarizes the articulatory properties of these rounded monophthongs based on Rögnaldsson's (2013) analysis of backness and height and Roger's (2014) analysis of lip posture.

<table>
<thead>
<tr>
<th>Articulatory property</th>
<th>[ʏ]</th>
<th>[œ]</th>
<th>[u]</th>
<th>[ɔ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backness</td>
<td>near-front</td>
<td>central-front</td>
<td>near-back</td>
<td>back</td>
</tr>
<tr>
<td>Height</td>
<td>mid</td>
<td>low-mid</td>
<td>high</td>
<td>mid</td>
</tr>
<tr>
<td>Lip posture</td>
<td>compressed</td>
<td>compressed</td>
<td>protruded</td>
<td>protruded</td>
</tr>
</tbody>
</table>

Table 1: Summary of Icelandic rounded vowels and their articulatory features
(Rögnvaldsson, 2013).

All of the vowels are pronounced with rounded lips, but they differ in backness, height and lip posture. The back vowels [u] and [ɔ] are both pronounced with protruded lips, but the front vowels [ʏ] and [œ] do not have protruded rounding. According to Rogers (2014), this is a cross-linguistic tendency; back rounded vowels tend to be protruded whilst front rounded vowels to be compressed.

For the prospective study, I chose to use the two vowel sounds, [u] and [ɔ] for the two rounded labels. Although they differ in height, they are both back vowels and limit backness as a confound. More importantly, they provide a good contrast in lip posture. Although both sounds involve labial protrusion, the vowel [u] is pronounced with a small mouth opening and tightly rounded lips, whereas [ɔ] is pronounced with a more open mouth and a lower jaw position (see Figure 2) (Mannell, 2009).
Moreover, the two vowels also differ in tongue placement (see Figure 3). Both vowels are back vowels, i.e. the tongue is positioned far towards the back of the mouth. However, [u] is raised and [ɔ] is retracted, i.e. the body of the tongue is raised toward the soft palate for [u] and pulled back into the pharynx for [ɔ]. In addition, they are of different height, [u] is high and [ɔ] is mid (Jones, 1972).

The stimuli consist of the two rounded labels along with an unrounded label (see Table 2). The set includes the label *boba* with the rounded vowel [ɔ], the label *bůba*...
with the rounded vowel [u], and the label *biba* with the unrounded vowel [i]. They are of the form CVCV (consonant-vowel-consonant-vowel) in which the first vowel is the critical vowel. Since stress falls on the first syllable in Icelandic and stressed vowels are long before a single consonant, the critical vowel is stressed and long for all three labels (Haugen, 1958).

<table>
<thead>
<tr>
<th>Rounded</th>
<th>Rounded</th>
<th>Unrounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>boba [ɔ]</td>
<td>būba [u]</td>
<td>biba [t]</td>
</tr>
</tbody>
</table>

*Table 2: Pseudoword stimuli.*

I use the same frame for all three labels to avoid confounding consonants or word length as in as in the *bouba/kiki* labels (Ramachandran & Hubbard, 2001). The relevant vowel is the only difference between the three words, ensuring that any possible sound-symbolic effects from the other sounds does not interfere with the results. I chose the consonant [b] for the frame because it produces relatively plausible Icelandic nouns that do not exist in the language already, and the word-final vowel [a] because it is a common female noun ending.

### 6.4 Procedure

The procedure and methods of the experiment are based on the previously discussed study by Ković et al. (2010). I aim to enlist 30-40 native speakers of Icelandic to participate. They will be 18-35 years old with normal or corrected-to-normal vision.

The experiment consists of two types of sound-symbolically congruent and incongruent conditions (see Table 3). Each condition includes two labels, one rounded label *boba/būba* and the unrounded label *biba*. In the congruent conditions, *boba* and *būba* are the correct labels for round figures and *biba* is the correct label for non-round figures. In the incongruent conditions, *boba* and *būba* correspond to non-round figures.
and *biba* to round figures.

<table>
<thead>
<tr>
<th>Conditions</th>
<th><em>boba</em></th>
<th><em>búba</em></th>
<th><em>biba</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent</td>
<td>round</td>
<td>-</td>
<td>non-round</td>
</tr>
<tr>
<td>Incongruent</td>
<td>non-round</td>
<td>-</td>
<td>round</td>
</tr>
<tr>
<td>Congruent</td>
<td>-</td>
<td>round</td>
<td>non-round</td>
</tr>
<tr>
<td>Incongruent</td>
<td>-</td>
<td>non-round</td>
<td>round</td>
</tr>
</tbody>
</table>

*Table 3: Experimental conditions.*

The labels will be presented to participants auditorily as feedback phrases in the experimental task ("Það var boba/búba/biba!" meaning 'It was a boba/búba/biba!). They will be recorded by a native speaker of Icelandic that is unfamiliar with the purpose of the experiment.

Following the experimental design of Ković et al. (2010), the experiment consists of a learning phase and a test phase. Beginning with the learning phase, participants will go through 9 trials to learn to match pseudowords with visual stimuli. In each trial a pseudoword (*boba*, *búba* or *biba*) will appear on the screen along with a round figure and an angular figure. Participants are asked to match the word to the figure they think it most accurately represents. To do this they press a left/right button on a keyboard corresponding to the vertical position of the visual stimuli. If they correctly match the word, they hear a pleasant chiming sound followed by a feedback phrase with the correct label ("Það var boba/búba/biba!"). If they incorrectly match the word, they hear an unpleasant buzzing sound and a feedback phrase with the correct label. The learning phase will end after either 21 trials or 9 correct responses in a row. In a subsequent test phase, participants will be presented with a label-object pairing and asked to press a button to indicate whether the pairing is a match or a mismatch.
In both the learning phase and test phase, response times will be recorded as a behavioral measure. Ković et al. (2010) found that participants that learned to map the pseudoword ‘mot’ to round shapes and ‘riff’ to spiky shapes were significantly faster to give correct responses, compared to participants who learned the reversed mapping. This difference in reaction time was consistent in both the congruent and incongruent condition for both match and mismatch trials. Ković et al. concluded that participants seemed to have a significant bias for a sound-symbolically congruent association. Based on their results, I will use response time as the main experimental measure.

7 Conclusion

I began this paper by presenting the received view that arbitrariness is a general property of language and sound symbolism is insignificant. I then argued that we should move beyond debate because the two systems are both prevalent in language and do not preclude each other. While arbitrariness has been considered a general property of language since the times of Saussure (1916/1966) and Hockett (1959), sound symbolism is increasingly being accepted in the same way (Lockwood, 2017).

I examined experimental research on cross-modal correspondences relevant to my paper and study proposal. These were sound-size correspondences and sound-shape correspondences. Although findings are not conclusive, both acoustic and articulatory properties have been reasoned as the main mechanisms behind sound-symbolic effects. I presented evidence from Ohtake and Haryu (2013), supporting the theory that acoustic properties are essential, and related Oda's (2000) experiment showing potential effects for an articulatory connection.

Based on the theory that vowel-shape correspondences are grounded in the articulatory properties of vowels, I proposed a study to assess the associations between rounded vowels and round shapes. I based the experimental design on a categorization
task by Ković et al. (2010), but used new pseudoword labels to contrast individual rounded vowels for speakers of Icelandic. I predicted that the pseudoword label búba would be more round-sounding and result in faster response times compared with the label bobo because of the differences in lip posture. The aim of the study is to determine whether there is a significant difference in effect between the two rounded vowels [u] and [ɔ]. If a difference is found, it could indicate that roundness associations are not a binary distinction of rounded vowels, but a graded effect that varies in strength depending on articulation. The prospective study is meant as a step in the direction of more accurately assessing vowel-roundness associations.
8 References


