TALKING SENSE

THE BEHAVIOURAL AND NEURAL CORRELATES OF SOUND SYMBOLISM

GWILYM LOCKWOOD

Max Planck Institute
for Psycholinguistics
Series

Talking Sense

The behavioural and neural correlates of sound symbolism

ISBN: 978-90-76203-83-6
Cover: *fuwafuwa* or *nurunuru*?

Cover image freely downloaded from http://paper4pc.com/bernese-mountain-dog.html in accordance with the site's copyright policy which allows anybody to use, store, display, reproduce, distribute and modify uploaded images.

Printed and bound by Ipskamp Drukkers b.v

Het onderzoek in dit proefschrift werd mogelijk gemaakt door een beurs van de International Max Planck Research School for Language Sciences, gefinancierd door de Max Planck Society, München, Duitsland.

Talking Sense
The behavioural and neural correlates of sound symbolism

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op

woensdag 12 april 2017
om 14.30 uur precies

door

Gwilym Lockwood

geboren op 21 september 1989
te Aylesbury, Verenigd Koninkrijk

Promotor:
Prof. dr. Peter Hagoort

Copromotor:
Dr. Mark Dingemanse

Manuscriptcommissie:
Prof. dr. Stephen Levinson
Prof. dr. Sotaro Kita (Warwick University, Verenigd Koninkrijk)
Dr. Suzy Styles (Nanyang Technological University, Singapore)

Table of contents

Chapter 1

Introduction to the thesis

Many PhDs start with a quote from a scientist/author/bearded man from centuries ago. This well-chosen quote demonstrates both the timeless nature of the important topic in question and precisely how far the field has come since then.

This isn't one of those PhDs.

I'm going to start this PhD with a cartoon from a book called Wicked Words (Deary, 1996), which I read when I was about eight years old:

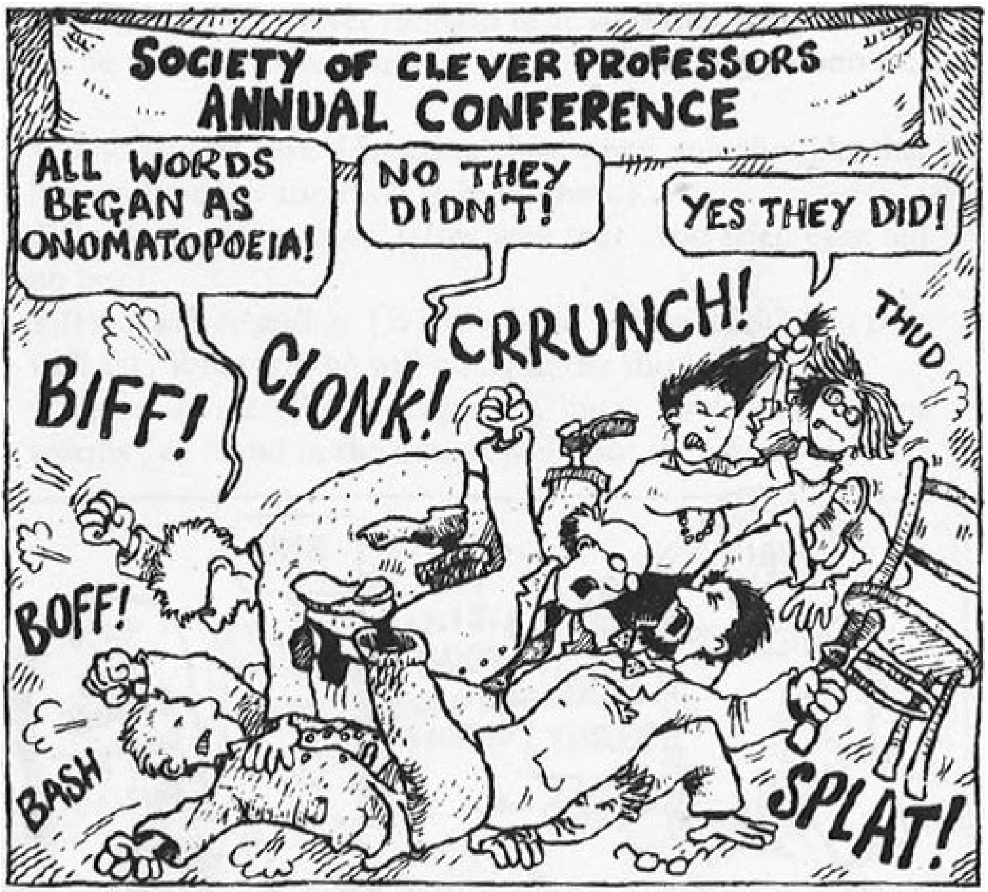


Figure 1: illustration of EvoLang from Wicked Words, children's book and thesis inspiration

I like this cartoon a lot. As a child, it was a fun way of introducing me to thinking about language and languages, and I read this book from the start to the end several times. Better still, it introduced me to phonaesthemes in a way that allowed me to insult people more expressively on the playground:

Figure 2: phonaesthemia for children

Now that I'm writing up three years of work on sound symbolism, I can see that the onomatopoeia cartoon not only encapsulates both the theories and the intensity around the evolution of language, it also illustrates our approach to onomatopoeia in general: they're sound effects and overdramatics, they're childish, they're limited in scope, they're the "playthings of language" (Müller, 1899). This marginalisation is understandable if we focus only on a handful of European languages, where onomatopoeia does tend to be a rather unserious, marginal part of language use, but for a long time, linguists figured that these playthings were all that we've got. Newmeyer wrote that the influence of iconicity was "vanishingly small" (1992). As most iconicity researchers begin their articles with this quote, those seem like famous last words.

But the thing is, onomatopoeia goes so much further than sound effects in comics, than animal noises for kids. Sound symbolism, or the resemblance-based mapping between aspects of form and meaning in spoken language (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015), is found across the world. Many languages have classes of ideophones, or marked words which depict sensory imagery (Dingemanse, 2012). Ideophones are often best translated into English and Dutch as adjectives or adverbs, but their role is often rather fluid, and ideophones are best thought of as a separate category of word. Moreover, ideophones are not an all-or-nothing kind of thing; languages with ideophones also have regular, arbitrary adjectives, adverbs, and verbs just as European languages do. Ideophones are used in addition to, rather than instead of, arbitrary words.

Despite ideophones being perfect examples of sound-symbolic words in natural language, they have been underused in experimental sound symbolism research, generally in favour of deliberately constructed pseudowords which exaggerated the sound-symbolic contrasts. The goals of this thesis are threefold; to bring ideophones into mainstream experimental psycholinguistics, to extend the field's knowledge of the neural signatures of sound symbolism, and to investigate the effect of individual differences in sound symbolism. Since ideophones tend not to be found in European languages, it can be a laborious and daunting task to create workable stimuli sets. All the materials used in this thesis are open for anybody to download and use in future experiments, which will hopefully go some way towards addressing the first goal of this thesis.

I used various experimental approaches for investigating what makes ideophones (and, later, sound-symbolic pseudowords) special. The main paradigm in this thesis is a learning paradigm, where Dutch speakers learned Japanese ideophones or arbitrary words (i.e. Japanese words which are neither ideophones nor sound-symbolic, but convey similar meanings). They saw the Dutch meaning, then they heard the Japanese word, then saw both written on a screen. Immediately afterwards, they performed a test, where they again saw the Dutch meaning then heard the Japanese word, and then they had to press a button to say whether or not that was the correct word pair. However, there was a manipulation that they weren't aware of; half the Japanese words were presented with their real Dutch meaning, but half the Japanese words were presented with their opposite Dutch meaning. For example, in the real condition, participants would learn that the ideophone *bukubuku* ("fat") meant *dik* ("fat"), whereas in the opposite condition, participants would learn that *bukubuku* meant *dun* ("thin"). The idea is that, if the sound symbolism in ideophones is identifiable and important in word learning, then it should be harder to learn ideophones with their opposite meanings than ideophones with their real meanings, but that there should be no difference between conditions with arbitrary words.

Rating paradigms with ideophones (and sound-symbolic pseudowords) were useful for assessing how iconic people find them. Participants saw a Dutch word, then heard an ideophone, and had to rate how well they think the Japanese word describes or sounds like what the Dutch word means. This was done on a scale of 1 to 7, where 7 meant that the Japanese word really sounded like the Dutch meaning, 1 meant that the Japanese didn't sound like the Dutch meaning at all, and 4 was for a neutral feeling that the Japanese word didn't describe the Dutch meaning particularly well or badly. This gave a good measure of how iconic a particular ideophone is. I also used two-alternative forced choice (2AFC) tasks in many experiments, where participants heard a Japanese ideophone and had to choose which of two antonymic Dutch translations suited the ideophone best. 2AFC tasks are a handy measure for quickly assessing how sensitive to iconicity an individual participant is, but the enforced dichotomy of it can also overstate the strength of an association.

In one chapter, I used a sentence evaluation task to distract the participants from the experimental manipulation. I changed the verb at the end of the sentence so that half the sentences were perfectly normal (e.g. *Hanako speaks French fluentl*y) while half the sentences were semantically impossible (**Hanako cooks French fluently*). Japanese speakers read a whole sentence, one word on the screen at a time, and had to judge whether or not the sentence made sense. By focusing on evaluating the sentences, they didn't notice that the real manipulation was whether the sentence used an ideophone or an arbitrary word to express the same meaning (e.g. *perapera* (ideophone) or *ryuuchouni* (arbitrary) for *fluently* in *Hanako speaks French fluently*).

Finally, in some of the experiments, I used electroencephalography (EEG) to measure participants' brain activity. By averaging many individual trials together into event-related potentials (ERPs), I could compare differences in how participants process two conditions, as well as looking at individual differences across participants. ERPs tell you *when*, rather than *where*, something happens in the brain. By looking at the time course of brain activity after an event, ERPs can distinguish what kind of cognitive process is implicated in the experimental manipulation.

Chapter 2 gives an overview of the experimental research into sound symbolism so far. Originally published as a review paper under the title "Iconicity in the lab: behavioural, developmental, and neuroimaging research into sound symbolism" (Lockwood & Dingemanse, 2015), it outlines how iconicity research spent a long history on the lunacy fringe of linguistics and psychology but how kiki-bouba experiments have been all the rage since 2001. This chapter calls for the field to embrace interdisciplinary approaches to iconicity, and to move on from mere observation of effects to attempts to explain the effects beyond just saying "it's synaesthetic".

In Chapter 3, I investigate the processing of Japanese ideophones in comparison to regular arbitrary Japanese adjectives or adverbs in a sentence reading task. Japanese native speakers read a whole sentence word by word and had to judge whether the sentence made sense or not. The actual experimental manipulation is about whether the Japanese word for *fluently* was an ideophone or not. I recorded participants' EEG during sentence reading to investigate how the brain responds differently to ideophones and arbitrary words. This paper was published under the title "Ideophones in Japanese modulate the P2 and late positive complex responses" (Lockwood & Tuomainen, 2015).

In Chapter 4, I move on to testing Dutch speakers. I investigate how Dutch speakers with no knowledge of Japanese learned Japanese ideophones and arbitrary adjectives. Participants learned Japanese words with either their real Dutch translation or their opposite Dutch translation. If sound symbolism in ideophones is recognisable and exploitable in word learning, this should result in a boost for the real condition (or hindrance for the opposite condition) for the group learning ideophones, but not for the group learning arbitrary adjectives. This paper was published under the title "Sound symbolism boosts novel word learning" (Lockwood, Dingemanse, & Hagoort, 2016).

In Chapter 5, I re-run the experiment from Chapter 4 while also testing participants' EEG. Replicating the behavioural results of Chapter 4 provides stronger evidence that the effect is real, and the use of EEG allows further investigation of the neural processes underlying the behavioural effect. This paper was published under the title "How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping" (Lockwood, Hagoort, & Dingemanse, 2016a).

In Chapter 6, I move on to looking at pseudowords. I created a set of more specific synthesised size/sound sound-symbolic stimuli so that I could investigate the real vs. opposite translation learning effect from Chapters 4 and 5 in a more graded way, with matching, neutral (i.e. neither matching nor mismatching), and mismatching conditions. These findings were published as a conference proceedings paper from CogSci (Lockwood, Hagoort, & Dingemanse, 2016b). I then re-run the same experiment with twice the original sample size.

In Chapter 7, I investigate individual differences in participants across a variety of sound symbolism rating tasks. To the best of my knowledge, all previous sound symbolism research papers involved participants doing one particular sound-symbolic task. There's nothing wrong with this; it builds up a pixel-by-pixel image of sound symbolism in general. But what if sound symbolism isn't a homogenous thing? What if participants do different sound-symbolic tasks differently, and their performance in one isn't related to the other? So, in this chapter, participants rate different sets of sound-symbolic stimuli — the ideophones and their real and opposite translations from Chapters 4 and 5, the synthesised size/sound pseudowords from Chapter 6, and some shape/sound stimuli adapted from Drijvers et al. (2015). I also use EEG for an ERP analysis.

These chapters are brought together and discussed in more detail in the final discussion section.

References

Deary, T. (1996). *Wicked Words*. UK: Scholastic.

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, *6*(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, *19*(10), 603–615. http://doi.org/10.1016/j.tics.2015.07.013

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-symbolism is disrupted in dyslexia: implications for the role of cross-modal abstraction processes. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Language Sciences*, 1246. http://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-Symbolism Boosts Novel Word Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. http://doi.org/10.1037/xlm0000235

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016a). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016b). Synthesized Size-Sound Sound Symbolism. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society [CogSci 2016]*. Philadelphia, PA: Cognitive Science Society.

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Language Sciences*, 933. http://doi.org/10.3389/fpsyg.2015.00933

Müller, M. (1899). *The Science of Language: founded on lectures delivered at the Royal Institution in 1861 and 1863.* (2nd edition). London: Longmans, Green, and Co.

Newmeyer, F. J. (1992). Iconicity and Generative Grammar. *Language*, *68*(4), 756. http://doi.org/10.2307/416852

Chapter 2

Iconicity in the lab: a review of behavioural, developmental, and neuroimaging research into sound symbolism.

Published as:
Lockwood & Dingemanse (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound symbolism. *Frontiers in Psychology*, 1246.

Abstract

This review covers experimental approaches to sound symbolism — from infants to adults, and from Sapir's foundational studies to 21st century product naming. It synthesises recent behavioural, developmental, and neuroimaging work into a systematic overview of the cross-modal correspondences that underpin iconic links between form and meaning. It also identifies open questions and opportunities, showing how the future course of experimental iconicity research can benefit from an integrated interdisciplinary perspective. Combining insights from psychology and neuroscience with evidence from natural languages provides us with opportunities for the experimental investigation of the role of sound symbolism in language learning, language processing, and communication. The review finishes by describing how hypothesis-testing and model-building will help contribute to a cumulative science of sound symbolism in human language.

Introduction

Despite the increasing acceptance and popularity of sound symbolism research in recent years, many articles about sound symbolism begin by defining it in opposition to arbitrariness. The traditional Saussurian (1916) or Hockettian (1959) view of language is outlined, the strengths of arbitrariness as a productive and compositional system (Monaghan & Christiansen, 2006) are described, the psychological and neuroscientific models of language which are built around arbitrariness (Friederici, 2002; Hagoort, 2013; Hickok & Poeppel, 2007; Levelt, Roelofs, & Meyer, 1999) are enumerated… and with a flourish, the latest sound symbolism research is uncovered to the reader. All is not what it seems!

This approach is certainly not without its uses; even relatively recently, the extent of sound symbolism within any given language was dismissed as "vanishingly small" (Newmeyer, 1992), and so the prerogative of sound symbolism researchers to point out the shortcomings and blind spots of a strictly arbitrary approach to language is understandable. However, to continue to present sound symbolism as an opponent to arbitrariness, rather than simply the opposite of arbitrariness, is unhelpful. The two systems are clearly happy enough to co-exist within language; with iconic links between sound/sign and meaning increasingly being accepted as a general property of language (Perniss, Thompson, & Vigliocco, 2010; Perniss & Vigliocco, 2014), it is time for a more constructive perspective.

Despite the fast growing interest in iconicity in general (as witnessed for instance in studies of sign language and gesture), there is still a relative dearth of experimental research on sound symbolism, especially when compared with the amount of psycholinguistic research based on arbitrariness. However, research into sound symbolism has been steadfastly gaining influence in fields like linguistics, psycholinguistics and cognitive neuroscience, opening up new opportunities for theoretical and empirical progress. What is needed now is a perspective that unites these bodies of evidence and shows where they converge or diverge. This review article brings together experimental findings from a wide range of fields — from behavioural experiments to developmental work and neuroimaging studies — and shows that there is now an exciting opportunity to develop a holistic account of the communicative functions and causal mechanisms of sound symbolism.

Definitions and history

The discussion of arbitrariness *versus* sound symbolism is nothing new. Plato's *Cratylus* describes a debate between Cratylus and Hermogenes about the origin of names, with Cratylus arguing that names are meaningful in themselves and by nature, and Hermogenes arguing that names are merely signifiers[1]. Socrates, the umpire of

---

[1] However, this entire debate was not conducted out of academic curiosity; rather, Cratylus had told Hermogenes that Hermogenes was not his real, natural name, assigned as it was by

the debate, acknowledges both points; he presents a Hamano-esque (1998) description of the "imitative significance of primary sounds corresponding to single letters of the alphabet", followed by the argument any name, even if it is natural, cannot perfectly describe its referent and thus some degree of linguistic convention is inherent to all names (Sedley, 2003)[2].

Arbitrariness and iconicity, "the source of more trouble than any other aspect of communicative behaviour" (Hockett, 1959), continued to set themselves apart throughout the Middle Ages and well into the 20th century. It was only in the middle of the 20th century that arbitrariness was fully enshrined as the principle cornerstone of language, basing linguistic theory upon Saussure's (1959) posthumously translated and published work on the arbitrariness of the sign and Hockett's (1959) assertion that arbitrariness is one of seven — later updated to thirteen (Hockett, 1960) — key design features of human language.

Competing motivations for arbitrariness and sound symbolism

The strength of arbitrariness was identified as the ability to combine symbols into limitless conventional forms, giving language far more communicative power in terms of the range of concepts and relations it can express, while also explaining why different languages have different forms for the same concepts. Crucially though, Hockett also acknowledged that while the design feature arbitrariness gives limitless possibilities to communication, it also "has the disadvantage of being arbitrary" (Hockett, 1960). This is a caveat with implications for learning and communication which has not always been addressed. Indeed, more recent studies have indicated that sound symbolism and arbitrariness mutually pick up each other's slack. Non-arbitrary form-to-meaning relationships facilitate learning as they are grounded in existing perceptual and cognitive systems (Cuskley & Kirby, 2013) and enable the grouping of similar words into categories (Farmer, Christiansen, & Monaghan, 2006). Arbitrariness facilitates the learning of specific word meanings (Monaghan, Christiansen, & Fitneva, 2011) and prevents the confusion of concepts which are similar but critically different (such as two almost identical mushrooms; one edible, one poisonous) (Corballis, 2002).

A system based solely on arbitrariness would pose immense learning difficulties, with no link between linguistic form and human experience, and would make communication less direct and vivid; a system based solely on sound symbolism would prevent specificity of communication because it can only offer limited

---

his parents. Thus provoked, Hermogenes became the first documented proponent of arbitrariness, arguing that any given group of people can determine their own labels for concepts and if he calls himself Hermogenes then he has the absolute right to do so since it is only a label for the person he is, thank you very much.

[2] Socrates concludes the debate by saying that it is far better to study the things themselves rather than their names, a suggestion which is somewhat less useful for models of language.

conceptual distinctions (Bühler, 1990). The recognition that sound symbolism and arbitrariness coexist in language is echoed in recent theoretical syntheses of arbitrariness and iconicity (Perniss et al., 2010; Perniss & Vigliocco, 2014). They can coexist because each brings its own advantages for learning words and using them in communication. By supplying perceptual analogies for vivid communication, sound symbolism allows for communication to be *effective*; by providing the lexicon with greater depth and distinction, arbitrariness allows for the *efficient* communication of concepts. The two systems lend themselves better to different communicative uses, which do not preclude each other, and are in fact complementary. The research is slowly leading the field towards a complementary view of language which features both sound symbolism and arbitrariness, but there are a few obstacles in the way, not least coming up with a widely-accepted and consistently-applied understanding of exactly what sound symbolism actually is.

Types of sound symbolism

While arbitrariness is defined by the absolute lack of relation between form and meaning, defining sound symbolism is somewhat harder; the sheer variety of depth and type of links between form and meaning, both within and across languages, means that there is no simple opposite of arbitrariness. Perniss et al. (2010) and Schmidtke et al. (2014) cover the various subtypes of sound symbolism in detail; a quick overview will be given here. The term *iconicity* is the closest cover-all term for communicative signs showing a resemblance between form and meaning, used as "a blanket term for a broad range of phenomena, including what has been referred to in the literature as sound symbolism, mimetics, ideophones, and iconicity" (Perniss et al., 2010). Iconicity can be applied to communication in visual, spoken, and other modalities, , can be manifested at all levels from phonetics to discourse, and is perhaps even present in animal communication (Hockett, 1959)..

In this review paper, we use the term *sound symbolism* to refer to iconicity in spoken language. Hinton et al. (Hinton, Nichols, & Ohala, 1994, 2006) define sound symbolism as "the direct linkage between sound and meaning", and divide it into *corporeal*, *imitative*, *conventional*, and *synaesthetic* sound symbolism. Cuskley and Kirby (2013) refine the latter two into *conventional* and *sensory* sound symbolism. Conventional sound symbolism is the regular correlation between specific sounds or clusters and specific meanings (such as with phonaesthemes). Conventional sound symbolism can also cover the correlation between sounds and grammatical categories, which is broadly equivalent to what Monaghan et al. (2011; 2014) call *systematicity*. This definition of conventional sound symbolism has a wider scope than most, as it goes further than Hinton et al., who don't consider sound symbolism as extending to grammatical categories, while Monaghan et al. also consider systematicity to be separate from sound symbolism, which they limit to phonaesthemes and sensory sound symbolism. Sensory sound symbolism is a natural connection where the word's form imitates aspects of the referent within or across modalities, and this imitation is often obvious across languages.

This classification echoes the description of sound systems outlined by Humboldt in 1836. "Since words always correspond to concepts, it is natural for related concepts to be designated by related sounds" (1836). Humboldt lists three ways in which sounds designate concepts: *direct imitation*, which broadly follows imitative sound symbolism or onomatopoeia; *symbolic designation*, whereby sounds "partly in themselves and partly by comparison with others produce for the ear an impression similar to that of the object upon the soul", and which most closely resembles sensory sound symbolism with the acknowledgement of some degree of conventionalism; and *analogical designation*, whereby "words whose meanings lie close to one another are likewise accorded similar sounds; but … there is no regard here to the character inherent in these sounds themselves", which most closely resembles conventional sound symbolism driven by statistical association, or systematicity. A closely related distinction is Gasser et al.'s (2010) two-way classification of iconicity as *absolute* or *relative*. *Absolute* iconicity is where there is a direct relation between form and meaning (as in onomatopoeic words for animal sounds). *Relative* iconicity is where related forms are associated with related meanings, as when a contrast between the vowels [i:a] depicts an analogous contrast in magnitude.

Many different terms and definitions have been used for sound-symbolic words, but *ideophone* is now the most widely used and accepted (Voeltz & Kilian-Hatz, 2001). Nuckolls (1999) defines ideophones as “lexicalised sound-imitative words”, while Dingemanse (2012) provides a more specific definition of ideophones as “marked words which depict sensory imagery”. Ideophones typically exhibit sensory sound symbolism, although there is always some degree of conventionalisation involved as well. Thus the Japanese ideophone *kirakira* ‘glittering’ shows sensory sound symbolism in that reduplication in the word is associated with a continuous meaning and the vowel [i] is associated with brightness, but it also has conventionalised aspects in that not all aspects of its meaning can be deduced from its sounds.

Sound symbolism is not confined solely to ideophones; in fact, the majority of sound symbolism research has focused on the cross-modal relations between individual sounds and sensory meanings, such as vowels and object size. There are also sound-symbolic links between certain combinations of sounds and meanings. Phonaesthemes are “frequently recurring sound-meaning pairings that are not clearly contrastive morphemes” (Bergen, 2004), such as such as *tw-* in English words like *twist*, *tweak*, *twizzle*, *twirl*, and *twine*. They show a mix of conventional and sensory sound symbolism (Kwon & Round, 2014), and are thought to be drivers of neologisms in language (Malkiel, 1994). Again, Humboldt wrote of such conventionalised forms having "undoubtedly exerted a great and perhaps exclusive dominance on primitive word-designation … and the new increment is formed by analogy with what is already present" (1836). This philosophical legacy has posed the question of how sound symbolism constitutes and affects language; it is now the responsibility of modern experimental approaches to bring iconicity out of the wild and into the lab to resolve the argument between Cratylus and Hermogenes with evidence as well as reason.

Behavioural experiments

There is a long history of behavioural research on sound symbolism, most of which has investigated the mappings between consonant/vowel types and the size or shape of visual stimuli in variations on experiments performed by Sapir (1929), Newman (1933), and Köhler (1929). Half a century of Generativism saw sound symbolism research fall out of favour somewhat, but this approach was brought back into fashion around the turn of the century (Kita, 1997; Klamer, 2002; Ramachandran & Hubbard, 2001; Waugh, 1994), and described in detail in Perniss et al. (2010). To begin with, it was enough simply to show that certain sounds have some kind of effect; this was an important rediscovery which brought sound symbolism in from the cold and into the wider attention of the field. More recently, there have been several studies in the last few years which have attempted to tease apart the separate roles of vowels and consonants, either by testing participants with individual phonemes or with non-words. These studies have also examined the effect of specific sounds on various different modalities, including strength, light, and taste.

Forced choice tasks with non- words

The standard paradigm in behavioural sound symbolism experiments is the *bouba/kiki* paradigm. Originally developed by Köhler, participants see two shapes — one spiky and one round — and two non-words — *takete* and *maluma* (later adapted to *kiki* and *bouba* by Ramachandran and Hubbard (2001)). Participants are then asked to say which non-word goes with which shape. Participants generally map the round shape with the "round" non-words (*maluma/bouba*) and the spiky shape with the "spiky" word (*takete/kiki*). Despite the methodologically sparse descriptions in Ramachandran and Hubbard (2001), this effect appears to be strong and consistent, and is the most well-known result showing that the relation between sound and meaning is not entirely arbitrary. This paradigm, and most variations of it, is perhaps the most obvious example of sensory sound symbolism.

Building on the *bouba/kiki* paradigm, various experiments have found consistent effects with better-controlled stimuli. The paradigm is affected by altering both individual consonants and vowels, but not by mode of presentation, as the effect was consistent regardless of whether the stimuli were presented auditorily or visually (Nielsen & Rendall, 2011). Systematically altering the placement of consonants and vowels in novel words addressed the shortcomings of Ramachandran and Hubbard's (2001) study, where the 95% success rate was down to the obvious distinction created by the non-words and novel shapes which were deliberately designed to be as different as possible. A follow-up non-word/shape matching experiment revealed a learning bias towards sound symbolism, albeit a weak one (Nielsen & Rendall, 2012). Two groups of participants were investigated; one which had been implicitly taught a congruent sound-symbolic pattern (plosives and spiky shapes, sonorants and curvy shapes) and one which had been implicitly taught an incongruent sound-symbolic pattern (plosives and curvy shapes, sonorants and spiky shapes). The first group performed above chance in the matching task while the second group performed at chance level, which demonstrates a learning bias towards sound symbolism. In a novel word generation task (Nielsen & Rendall, 2013), participants were found to use both vowels and consonants to form sound-symbolic associations. Participants used sonorant consonants and rounded vowels for curvy *bouba* figures and plosive consonants and non-rounded vowels for spiky *kiki* figures. Participants also favoured vowels with relatively close articulation to the co-articulated consonant (such as a frontal [i] following the strident consonants [t] and [k] and the “frontal” consonants [m], and [n]) and showed a dispreference for combining consonants and vowels which were relatively further apart. This suggested once more that consonants trump vowels when it comes to non-word sound-symbolic perception of visual contours, but that both types of sound do have a role.

The *bouba/kiki* paradigm has also been informative about language in populations different from psychology undergraduate students participating for course credit (Henrich, Heine, & Norenzayan, 2010). A first cross-linguistic and cross-cultural replication of Köhler's (1947) *maluma/takete* paradigm was Davis’ (1961) study of English and Tanzanian children. More recently, Bremner et al. (2013) replicated the *bouba/kiki* paradigm with Himba participants in Namibia for sound-to-shape matching but not taste-to-shape matching. The Himba have no written language and very little exposure to Western culture, which is helpful in ruling out cultural or orthographic effects such as associations with brand names or associations with the shape of the letters (such as how the letter K is spikier than the letter O).

Finally, developmental disorders involving impaired cross-modal integration also affect participants' accuracy at the *bouba/kiki* paradigm. High functioning autistic participants were significantly worse than non-autistic participants at matching *kiki*-like words to spiky shapes and *bouba*-like words to curvy shapes, although they still categorised the stimuli at above-chance level; low functioning autistic participants performed at chance level (although this may be due to the nature of the task) (Occelli, Esposito, Venuti, Arduino, & Zampini, 2013). Occelli et al. theorise that this is linked to a global deficiency in multisensory integration in autistic people, suggesting that the cross-modal correspondence effect is linked to motor and sensory integrative processes in the left inferior frontal gyrus. Dyslexic Dutch speakers, meanwhile, perform above chance at *bouba/kiki* paradigms but worse than non-dyslexic Dutch speakers (Drijvers, Zaadnoordijk, & Dingemanse, 2015). This reinforces the claim that cross-modal abstraction is involved in making sound-symbolic links.

Task effects

The robustness of the *bouba/kiki* paradigm relies in part on the nature of forced choice. When it uses four target stimuli rather than two, participants are less successful at making congruent sound-symbolic matches (Aveyard, 2012). Moreover, the use of three rounds of testing showed that participants use different strategies depending on whether the paradigm is a two- or four-alternative forced choice task. When there were only two choices, participants used a consonantal sound-symbolic strategy instantly, and general accuracy for incongruent trials improved over three rounds of testing, indicating that participants were able to use separate strategies for congruent and incongruent trials after some experience. When the number of choices was increased to four, participants were less aware of the manipulation and were slower to incorporate consonantal sound symbolism into their decision making, although this did emerge by the third round. The main effect of linking sonorants to curviness and plosives to spikiness is in line with most behavioural research, but introduces some important variables which show how easily this sensitivity to consonantal sound symbolism can be affected by experimental set-up.

Moving beyond shape

While the *bouba/kiki* paradigm has been very popular for sound symbolism research into shape, other experimental approaches are more useful for investigating other sensory modalities. Hirata et al. (Hirata, Ukita, & Kita, 2011) found an effect of lightness on sound sensitivity. Participants were better able to identify consonants when they heard and saw congruent sound-light pairings (i.e. voiceless consonants with light visual stimuli, voiced consonants with dark visual stimuli) than incongruent sound-light pairings. However, there was no effect of consonant type when participants had to identify whether a visual stimulus was light or dark.

Links between sound and emotion have also been investigated, but these are more likely to rely on indexical interpretations of affective prosody rather than on iconicity in the sense of structural resemblance (Majid, 2012).

Most of the research presented so far has focused on the properties of consonants, but sensory sound symbolism with vowels is well-attested too, especially for size (Sapir 1929). Thompson and Estes (Thompson, 2013; Thompson & Estes, 2011) investigated sound symbolism and object size links by addressing the forced dichotomy of two-alternative forced choice matching in a slightly different way from Aveyard (2012). They showed five different sizes of novel object set against a picture of a cow as a point of comparison, and asked participants to choose the most appropriate name from a selection of three-syllable non-words which varied the number of small-sounding (such as [i]) and large-sounding (such as [a]) vowels. Participants chose non-words with increasing numbers of large phonemes for increasingly large objects, which shows that sound symbolism marks graded cross-modal mappings rather than just marking contrasts. Meanwhile, it appears that the evidence for an acoustic mechanism for sound symbolism is stronger than that for a kinaesthetic mechanism, a perennial debate which goes back to Sapir (1929) and Newman (1933). Ohtake and Haryu (2013) performed a series of experiments which separated acoustic features of vowels and the size of the oral cavity while asking participants to categorise the size of a visual object. Participants were faster to categorise object size when hearing the vowels [a] and [i] in congruent conditions, i.e. when [a] was presenting with a large object and [i] with a small object. However, there was no effect when participants categorised object size while holding objects in their mouths to simulate the oral cavity shape made when pronouncing the vowels [a] and [i]. This suggests that the main driver of the effect is the acoustic properties of the vowels, rather than their articulatory properties.

The acoustic properties of vowels have also been found to elicit cross-modal correspondences related to taste (Simner, Cuskley, & Kirby, 2010). Participants were given taste samples of four taste types — sweet, sour, bitter, and salty — and adjusted four sliders — F1, F2, voice discontinuity, and spectral balance — to create a vowel sound which best fit the taste. Participants consistently assigned lower F1 and F2 frequencies (approximating higher, more back vowels) to sweet flavours and higher F1 and F2 frequencies (approximating lower, more front vowels) to sour flavours, with salty and bitter flavours falling in between. They posit that these patterns may have influenced vocabulary construction for taste terminology. Interestingly, this spectrum doesn't quite fit along the same lines as most sound-symbolic vowel associations, which tend to run on a spectrum from [i] to [a] as illustrated in figure 1. Given that Anglophones find it especially hard to describe and discriminate between tastes and smells according to their properties (as opposed to their sources) when compared to other senses (Majid & Burenhult, 2014), perhaps it is to be expected that Anglophone participants may not map sounds onto tastes in the same way as other senses. It is also hard to say what kind of sound-symbolic links drive this effect. It is probably sensory sound symbolism, but there may be conventional aspects involved; the word *sour* is pronounced with a lower vowel than the word *sweet*, which mirrors the associations made by the participants.

Differences between back vowels and front vowels have been found in various studies. Cuskley (2013) investigated non-words and visual motion by asking participants to direct the motion of a ball to match a non-word. Participants made the ball travel more slowly in response to back vowels, and made the ball travel more quickly in response to front vowels and syllable reduplication with vowel alternation (the apophonic direction of vowel alternation in reduplicated syllables was not tested; forms such as *kigu* and *kugi* were treated as the same). However, whether this mapping is consistent is unclear; Thompson (2013) performed a similar study and found only a small and non-statistical trend towards assigning faster ratings to names containing front vowels.

Maglio et al. (2014) linked front vowels to conceptual precision with two studies on vision and concepts. Participants were asked to perform a geographical analysis of a fictional city. When the city's name featured more front vowels than back vowels, participants divided the city into smaller, more precise geographic regions, and vice versa, which Maglio et al. refer to as visual precision. Participants were also more precise when asked to describe the actions of a person when there was a front vowel association. They saw a person writing a list and were told that this person was performing a "sheeb task" or a "shoob task"; when asked to describe the person's behaviour, participants replied with conceptual precision about the action in the front vowel condition (e.g. "the person is writing a list" when performing the "sheeb task"), and replied with conceptual breadth about the action in the back vowel condition (e.g. "the person is getting organised" when performing the "shoob task"). This may actually be an indirect measure of the typical vowel-size correspondences, with the participants associating back vowels with size in general and then applying the size distinction to visual or conceptual precision. Maglio et al. then performed a series of experiments on high vs. low-level thought; these linked front vowels to low-level thought and back vowels to high-level thought. Back vowels in an ice-cream product name made people focus on how good it tastes rather than how easily accessible it is; back vowels in a skin lotion product name made people focus on how effective it is, rather than how attractive the packaging is; and back vowels in a back pain treatment made people focus on how long-lasting the pain relief is, rather than how arduous the procedure is. Maglio et al.'s research provides interesting evidence that specific vowel changes may elicit different mental representations. This probably examines conventional sound symbolism rather than sensory sound symbolism, as vowel size does not map onto literal sensory size but a more metaphorical magnitude of abstract concepts.

Some studies have linked cross-modal associations between linguistic stimuli and colour to synaesthesia. Moos et al. (2014) investigated vowel sound and colour associations in synaesthetes and control participants. They found that increased F2 (such as in front vowels like /i/) was associated with increased yellowness and greenness on the colour spectrum, while increased F1 (such as in open vowels like /ɑ/) was associated with increased redness. This was found in both synaesthetes and non-synaesthetes, although far more strongly in the synaesthetes, which suggests that grapheme-colour synaesthesia is at least partially based on acoustic properties of the sounds associated to the graphemes, and provides further evidence that synaesthesia may be an exaggeration of general cross-modal associations which most people have. Shin and Kim (2014) likewise investigated colour associations in synaesthetes by comparing the associations of Japanese, Korean, and English graphemes in trilingual synaesthetes. Despite the small sample size, they found that colour associations were broadly similar across participants and across languages for graphemes which expressed the same sounds, showing that grapheme-colour synaesthesia for individual graphemes is based on the sounds which the graphemes express. In experiments with synaesthetic Japanese speakers, Asano and Yokosawa (2011) found that consonants and vowels independently influence the colours which synaesthetes ascribe to the hiragana and katakana Japanese writing systems, and that this effect was not due to visual form. Their results show a tendency for front vowels and voiceless consonants to be associated with brighter colours, and for back vowels and voiced consonants to be associated with darker colours, which follows the general synaesthetic patterns set out by Marks (1978). The fact that most of the participants are synaesthetic in these three studies makes it hard to say which type of sound symbolism is under investigation here, but it is likely to be sensory sound symbolism.

Summary of attested cross-modal correspondences

Non-word behavioural experiments have been useful in establishing broadly consistent cross-modal associations between sound and other sensory modalities, and these seem to overlap with synaesthetic associations. When presenting full non-words, consonants seem to have greater prominence than vowels in terms of what participants perceive and how they formulate sound-symbolic strategies; however, both consonants and vowels do influence participants' judgements. Voiced consonants and low back vowels are consistently associated with roundness, darkness in colour, darkness in light intensity, and slowness (although in the case of voiced consonants, only by comparison with voiceless consonants). Voiceless consonants and high front vowels are consistently associated with spikiness, brightness in colour, brightness in light intensity, and quickness. Moreover, vowel height and size is linked with physical size, with low vowels and back vowels being linked to big objects and high vowels and front vowels being linked to small objects. Taste conflates the two acoustic properties of vowels; sweetness is linked with high back vowels and saltiness is linked with low front vowels. This is illustrated in Figures 1 and 2.

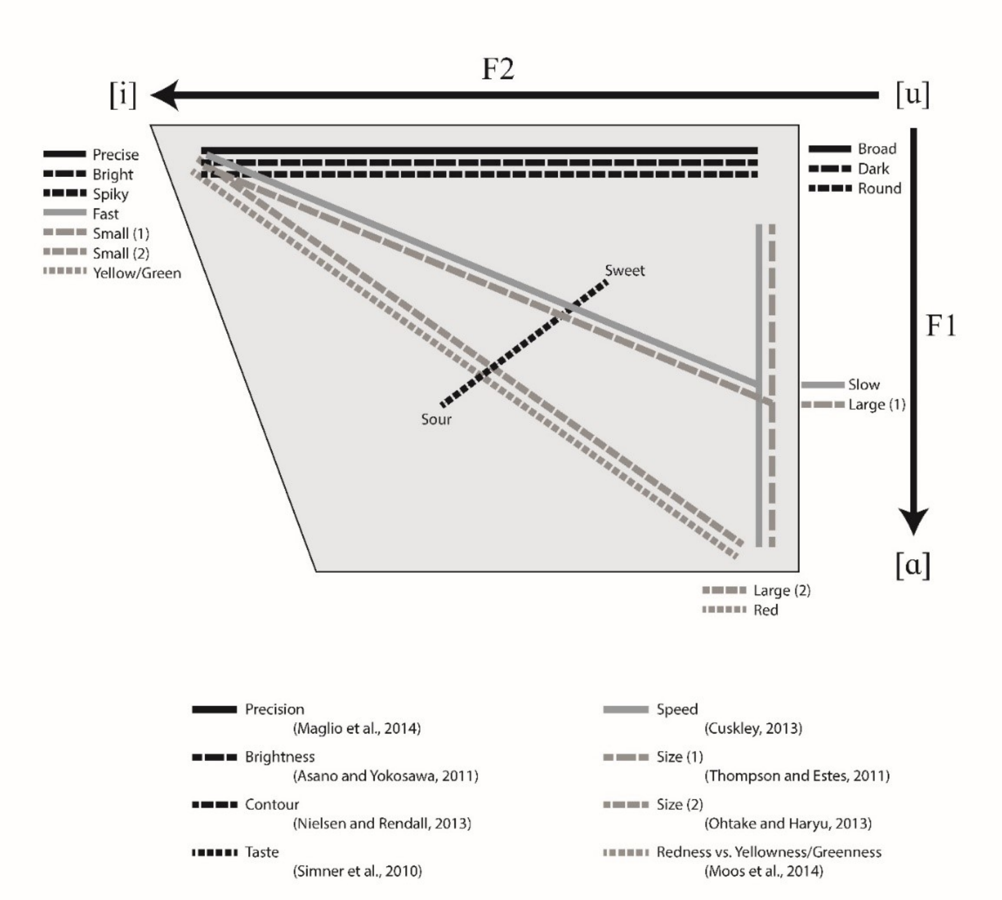


Figure 1: diagram of attested cross-modal mappings to linguistic sound represented on typical vowel space.

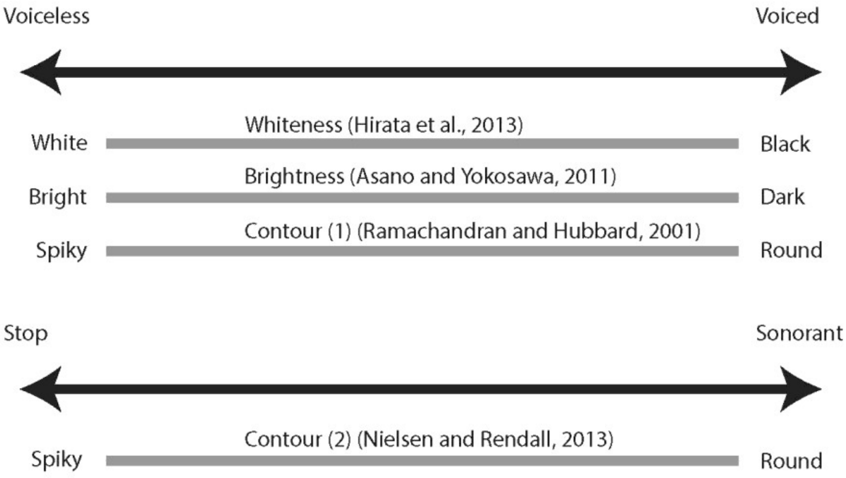


Figure 2: diagram of attested cross-modal mappings to linguistic sound for consonant properties

Moving beyond non-words

Despite the progress made with behavioural research on non-words, the insights it provides into language processing are limited. Non-word stimuli are carefully designed to provide maximal distinction between the sensory properties of the referent and the linguistic factors of interest, such as consonant voicing, vowel height and backness, and lip rounding. Not only does this introduce experimenter bias concerning which properties of language are sound-symbolic, it also means that the language stimuli used are not necessarily reflective of spoken language if such maximal distinctions do not occur naturally, and any existing findings may be an overstatement of the cross-modal associations that people make with real language. One way to address this problem is to use existing sound-symbolic words to address the question of how sound symbolism in natural language is (or isn't) associated with other sensory modalities; and among existing sound-symbolic words, ideophones are a prime source of information about sound-symbolic mappings.

Most experimental work on ideophones has been conducted using Japanese, which has an extensive, commonly-used and well-documented set of ideophones (Akita, 2009a; Hamano, 1998; Kita, 1997). Most studies have found that participants with no knowledge of Japanese perform significantly above chance at guessing the meaning of ideophones. Oda (2000) performed a series of forced choice tasks with Japanese ideophones on two groups of native English speakers. The first group heard a native Japanese speaker read out the ideophones and were asked to focus on the sound before performing the tasks. The second group heard a native Japanese speaker read out the ideophones and were then asked to pronounce the words themselves before performing the tasks. The two tasks were picking the correct ideophone out of three options for one English definition, and matching two minimal pair ideophones to the two English definitions, which were accompanied by illustrations of the texture or movement. Both groups could guess the meaning of the ideophones at an above chance level of accuracy, and this accuracy was modulated by articulation; the group which pronounced the words themselves were significantly better at matching unfamiliar ideophones to English definitions. In opposition to studies such as Ohtake and Haryu (2013), Oda's result suggests that articulation does play a role in establishing the form-meaning relationship of ideophones. The question over whether sound symbolism is driven by acoustic *or* articulatory mappings is perhaps too reductive; it seems that both mechanisms are involved depending on the nature of the task.

Iwasaki et al. (2007a) conducted similar experiments with Japanese pain vocabulary, and found that non-Japanese speakers could accurately categorise ideophones expressing pain according to the type of pain they express. However, Japanese sound symbolism is not always entirely transparent to other speakers. In another study, Iwasaki et al. (2007b) found that English speakers with no knowledge of Japanese could make accurate semantic judgements about ideophones which referred to specific sound qualities but the same speakers made very different semantic judgements about ideophones concerning beauty and pleasantness. It is unclear whether this is due to the fact that sound-to-sound mappings do not cross modalities and are therefore more transparent, whether these particular ideophones expressing beauty were just more on the conventional side of the continuum and therefore less obviously iconic, or due to cultural differences over what constitutes beauty.

Iwasaki et al. further found that English speakers were relatively better at categorising ideophones describing manners of laughter (e.g. giggling and chuckling according to semantic dimensions like pitch and gracefulness) than ideophones describing manners of walking (e.g. strolling and lumbering according to semantic dimensions like pace and steadiness). Iwasaki et al. attributed this to the same kind of vowel and consonant voicing contrasts which have been found in non-word studies, such as large vowels being linked with large strides and loud laughs. However, it also shows that ideophones are not completely intuitive to speakers of other languages and depend in some part on the specific semantic context provided by the experimental set-up. In a developmental study (Imai, Kita, Nagumo, & Okada, 2008), Imai et al. generated some novel ideophones for manners of motion based on Hamano's (1998) phonosemantic classification of Japanese ideophones, and Japanese adult participants completely agreed with the novel ideophones' intended meanings. This supports the idea that at least some of the sound-symbolic patterns in Japanese ideophones are sufficiently systematic enough to be productive (Oda, 2000; Yoshida, 2012). When naïve English speakers were tested with these novel ideophones, the intended meanings were still categorised at above chance level, thus confirming previous behavioural research on Japanese ideophones with novel forms. All of these studies with Japanese ideophones show that there is enough sensory sound symbolism in ideophones for speakers of other languages to be sensitive to the meanings, and that there may be additional conventional sound symbolism in ideophones which is more informative for native speakers.

The role of prosody

Similar above chance categorisation patterns have been found with ideophones in various languages, not just Japanese. Dingemanse et al. (2016) took ideophones from five languages across five semantic domains, and presented naïve participants with four versions of the stimuli in two-alternative forced choice tasks — the original ideophone recordings, a rich resynthesis using the original recordings' phoneme durations and prosody, a phoneme-only resynthesis and a prosody-only resynthesis. Ideophones in the original recordings and in the rich resynthesis condition were both categorised at above-chance accuracy, but ideophones in the phoneme-only and prosody-only resynthesis conditions were not. This indicates that both phonemes and prosody are important for cross-linguistic effects of iconicity. This finding is corroborated by evidence that around 80% of ideophones are given special prosodic attention and emphasis in natural speech — prosodically foregrounded (Dingemanse, 2013) — and that certain prosodic profiles in non-words have reliable semantic associations (Nygaard, Herold, & Namy, 2009).

Some non-ideophonic lexical words also show these effects. Kunihira (1971) conducted experiments using apparently arbitrary Japanese words in forced choice tests and found that English speakers were able to accurately categorise them, even though they were not ideophones. Responses were most accurate when the words were pronounced with "expressive voice", i.e. exaggerated prosody. This suggests sound-symbolic interpretations can be elicited even for arbitrary words — a viewpoint that reinforces the crucial role of expressive prosody. Nygaard et al. (2009) used Kunihira's stimuli in a learning task, and found that English speakers were quicker to learn and quicker to respond to Japanese words paired with correct English translations (e.g *hayai* and *fast*) than when paired with opposite (e.g. *slow*) or unrelated (e.g. *blunt*) English translations. Nygaard et al. stop short of linking particular sounds or properties of the words to particular meanings, instead suggesting that reliable sound-meaning mappings — regardless of whether this sound symbolism is sensory (i.e. presumably recognisable across languages) or conventionalised (i.e. recognisable only within a particular language) — "may constrain novel word learning and subsequent word retrieval and recognition by guiding processing to properties and meaning within a particular semantic context".

The same research group expanded the scope of this research to include antonym contrasts in ten different languages; monolingual English speakers allocated the antonyms correctly at above chance level in two-alternative forced choice testing, although consistency varied across individual items and may indicate the inherent probabilistic variability in the degree of sound symbolism in supposedly arbitrary words (Namy, DeFife, Mathur, & Nygaard, submitted for publication; Tzeng, Nygaard, & Namy, submitted for publication). These findings were partially replicated in a study comparing synaesthetes and non-synaesthetes, which found that both groups guessed certain meanings at above chance accuracy, and that the synaesthetes did so more strongly than the non-synaesthetes (Bankieris & Simner, 2015). However, there are two crucial caveats with these stimuli. Firstly, six of the ten languages used in these studies are rich in ideophones and poor in ordinary adjectives (Indonesian, Korean, Tamil, Mandarin, Turkish, and Yoruba), which means that this study may well have indirectly studied ideophones rather than arbitrary antonyms. Secondly, the four non-ideophonic languages (Dutch, Albanian, Gujurati, and Romanian) are all Indo-European; this means that they cannot be treated as independent because of potentially shared linguistic features, and moreover their meanings may be more transparent to native English speakers if they are cognates, especially in the case of Dutch and Romanian. Unfortunately, these studies are not yet publicly available (despite their crucial role in other published work), and so we cannot do more than speculate here.

Developmental experiments

While the extensive behavioural literature attests that sound symbolism has persistent and varied effects on language processing and use, a frequent criticism is that these patterns of association are conditioned because of orthographic influences; people might only consider the sound [b] to be rounder than the sound [k] because the letter *b* is rounder than the letter *k*. However, studies on early language development have shown that this is not the case. Studies with pre-literate children and young infants rule out such orthographic effects. Developmental experiments with infants also provide a different window into sound symbolism from learning experiments with adults. Experiments with infants examine existing cross-modal associations and how infants exploit these during early language development, whereas learning experiments with adults examine how sound symbolism affects memory, and are necessarily influenced by the adults' first language.

Mixed results for bouba/kiki paradigms

The *bouba/kiki* paradigm, with its sensory sound symbolism links, can be easily adapted for infants and young children, although results have been mixed. Ozturk et al. (2013) and Fort et al. (2013) tested 4-month old infants with preferential looking procedures, using fully reduplicated non-words with no word-internal vowel contrasts (e.g. *kiki*, *bubu*). Ozturk et al. presented one shape together with one auditory non-word and measured gaze duration, while Fort et al. presented two shapes side by side together with one auditory non-word and investigated whether infants preferred looking at a particular shape. The additional complexities of Fort et al's experimental set-up proved to be too much for the infants, as they found no preferential looking effects; they "tentatively argue that the complexity of their design might have masked the infants' emerging sound-symbolic matching abilities". However, Ozturk et al. found that infants looked for longer durations at shapes which were presented with incongruent non-words. Moreover, they found that this only happened for non-words where both vowels and consonants were typically sound-symbolic; the infants would match *bubu* with the curvy shape and *kiki* with the spiky shape, but would not make the same distinctions when comparing *kiki* and *kuku* or *bibi* and *bubu*. The adult control group, on the other hand, only needed either a vowel contrast or a consonant contrast to make cross-modal associations. When taken together, these results suggest that there is an effect of sound symbolism in infants, but that it needs both consonants and vowels to make the stimuli maximally distinct and that only very straightforward designs may detect the effect. This also appears to show that infants are less sensitive to sound-symbolic contrasts than adults are, which implies that increased exposure to language in fact increases sensitivity to sound-symbolic associations. This is supported by a study on pitch-size associations in 4- and 6-month old infants, which found that 6-month old infants make typical associations between pitch and size while 4-month old infants do not (Fernández-Prieto, Navarra, & Pons, 2015). The apparent conflict in results between Fort et al. and Ozturk et al. shows that iconicity may be strong enough for infants to detect, but not strong enough for this effect to persist through more complicated tasks.

Maurer et al. (2006) replicated Ramachandran and Hubbard's (2001) *bouba/kiki* results with 2.5 year old children, which ruled out orthography as a confound as these children could not yet read. Spector and Maurer (2013) developed this experiment with slightly updated stimuli, using fully reduplicated non-words with no word-internal vowel contrasts rather than the typical *bouba/kiki* words used in the previous study. The toddlers were presented with two visual shapes, and then asked by an adult to point to the non-word of interest (e.g. "can you point to the *koko*?"). As predicted, the toddlers associated curvy shapes with rounded vowels and spiky shapes with non-rounded vowels. One possible factor is the direct interaction with an adult experimenter rather than pre-recorded stimuli. Nygaard et al. (2009) have established that adults use exaggerated and semantically-predictable prosodic profiles when pronouncing non-words in child-directed speech, and this may have provided the kind of prosodic foregrounding which helps to identify ideophones in natural language.

There have also been several developmental studies on the acquisition and use of Japanese ideophones, which show that both Japanese and non-Japanese children are highly sensitive to the sound-symbolic properties of Japanese ideophones. Iwasaki et al. (2007b) cite Ishiguro (1993), who found that children create their own idiosyncratic ideophones before fully acquiring conventional ones, and that children acquire ideophones expressing sound before acquiring ideophones expressing motion, shape, psychological states, or other sensory modalities. This ties in with Iwasaki et al.'s and Oda's research, which showed that participants with no knowledge of Japanese were more accurate at categorising ideophones expressing sound, and confirms the prevalence of sensory sound symbolism in ideophones.

The sound-symbolic bootstrapping hypothesis

Imai et al. (Imai et al., 2008) created novel Japanese ideophonic motion verbs and tested them on Japanese and English-speaking adults (as described in the behavioural section). They then tested 25 month old Japanese children with a verb learning task, and found that the children could generalise the ideophonic verbs to new situations, but could not do the same for the non-sound-symbolic verbs. Imai et al. concluded that sound symbolism provides a scaffold on which children can map semantic and syntactic information. Echoing Gentner & Boroditsky's arguments (2001) that actions unfold over time and are impermanent whereas objects are stable, which is why children tend to focus on objects and tend to acquire nouns first, Imai et al. propose that the sound-symbolic scaffolding provided by the ideophonic verbs helps children to isolate the action and therefore facilitates verb learning. Kantartzis et al. (2011) replicated Imai et al.'s results in experiments with English children using the same novel verbs based on Japanese sound-symbolic patterns. This provided evidence towards a cross-linguistic – or, perhaps more accurately, language-independent – early sensitivity towards sound symbolism, and also shows that Japanese ideophones contain sensory sound symbolism and not just conventional sound symbolism. Kantartzis et al. also point out that it is unclear what exactly the English children recognise as sound-symbolic; it could be the phonetics, the phonotactics, the prosody, or a combination of all three.

Yoshida (2012) developed the paradigm further and carried out more extensive tests, and made several important points. Firstly, sound symbolism aided verb acquisition in Japanese and English children equally, despite the Japanese children's greater exposure to and familiarity with the Japanese mimetic-style novel verbs. Secondly, this equal language-independent sensitivity to sound symbolism exists despite the vast difference in general iconic input between Japanese (where parents make extensive use of ideophones to children) and English (where parents do use a lot of onomatopoeia to children, but they do so more idiosyncratically and less often than Japanese parents do). Thirdly, by including both novel verbs and novel actors in the task, she showed that the sound-symbolic scaffolding proposed by Imai, Kita et al. (Imai & Kita, 2014; Imai et al., 2008) helps children to isolate the action by excluding the identity of the actor, rather than just by focusing on the action. Yoshida proposes that infants are universally sensitive towards sound symbolism, but this sensitivity attenuates in adulthood as their native language's conventionalised forms dictate which possible forms of sound symbolism are acceptable; this mirrors the well-established pattern of infant sensitivity to cross-linguistic phonemic differences, which attenuates with age. The proposal of a sound-symbolic bootstrapping hypothesis is also supported by ideophone usage studies, which have shown that Japanese children as young as two years old use ideophonic verbs frequently and productively (Akita, 2009b) and that Japanese parents are five times more likely to use ideophones to children than they were to other adults when describing the same scene (Maguire et al., 2010). The finding that ideophones are more geared towards children initially appears to sit uncomfortably with the finding of Ozturk et al. (2013), which suggested that infants were less sensitive to sound symbolism than adults.

However, perhaps a reasonable middle ground is that children are more sensitive to sound symbolism as long as there are enough sources in the input to make associations from, while adults are less sensitive to sound symbolism in terms of forming associations but can form associations from a more limited input.

Finally, Laing's (2014) reanalysis of a longitudinal case study (Elsen, 1991) provides another example of how sound symbolism bootstraps language acquisition. Laing examined Elsen's detailed dataset of German infant Annalena and investigated the development and role of onomatopoeic forms. Annalena used onomatopoeic forms extensively, constituting almost 40% of her vocabulary at 11 months, but the relative proportion of onomatopoeia in Annalena's vocabulary tailed off with age. Annalena systematically replaced onomatopoeic forms with conventional words according to her phonological ability, meaning that onomatopoeic forms were retained longer when their conventional forms were phonologically more difficult. This shows how both sensory and conventional sound symbolism in infancy works alongside the developing lexicon and can bootstrap phonological development.

Neuroimaging experiments

Behavioural research into sound symbolism has been instrumental in telling us that there is a robust effect of sound symbolism on language tasks, and that this effect can be modulated by various different linguistic changes. However, neuroimaging research is needed to establish how the brain recognises, processes, and constructs sound symbolism. There has been far less neuroimaging research on sound symbolism than behavioural, but the handful of existing studies make interesting suggestions about sensory embodiment, synaesthesia, and multisensory integration.

ERP and fMRI evidence

Some neuroimaging experiments on ideophones have essentially used behavioural paradigms with simultaneous EEG recording to investigate ERPs. Kovic et al. (2010) conducted a novel word learning experiment, which established that participants were quicker to identify novel objects with congruent sound-symbolic non-word names than incongruent or arbitrary non-word names. They then tested two groups of participants; one group learned congruent sound-symbolic names for pointy and round objects (i.e. *shick* for a pointy object and *dom* for a round object), the other group learned incongruent sound-symbolic names (i.e. *shick* for a round object and *dom* for a pointy object). The experiment presented a name auditorily and then an object visually, and the participants had to decide whether the object and name matched. The first group were quicker to identify correct conditions and quicker to reject incorrect conditions than the second group, which corroborates other behavioural evidence that sensory sound-symbolic congruence has an object recognition facilitation effect. Moreover, objects with congruent sound-symbolic names elicited a stronger negative wave than incongruent ones in the 140-180ms window after the presentation of the object. This effect was observed at the occipital regions, home of the visual cortex, and Kovic et al. suggest that the early negativity represents auditory-visual integration during early sensory processing.

Arata et al. (2010) used the *bouba/kiki* paradigm on 12 month old infants, simultaneously presenting a shape and a non-word in congruent and incongruent conditions. The infants were found to be sensitive to sound-symbolic matches and mismatches, showing differentiated wave patterns across both conditions after 200ms post-stimulus. This may have been the P2, an ERP component which has been linked to phonological and semantic analysis. Arata et al. claim that their results support the claim that infants are synaesthetic or like synaesthetes (Maurer & Mondloch, 2004), potentially due to having more cortical connections than adults do, resulting in their ability to detect sound symbolism. Asano et al. (2015) performed a similar experiment on 11 month old infants, this time presenting the stimuli sequentially; the infants were first shown a spiky or curvy novel object, and then heard the non-word *kipi* or *moma*. This study found a later effect, with more negative ERPs in the 400-550ms window for incongruent stimuli compared to congruent stimuli. Asano et al. argue that infants use sensory sound-symbolic congruency to anchor novel sounds onto meaning, thus enabling them to establish that linguistic sounds have real world referents.

There are fewer neuroimaging experiments specifically aimed at revealing the brain locations associated with ideophone use and understanding, probably because of the relative lack of knowledge of ideophones outside the field of linguistics. However, a few neuroimaging studies using ideophones do exist. Osaka and his group conducted a series of fMRI studies (Osaka, 2009, 2011; Osaka et al., 2003; Osaka, Osaka, Morishita, Kondo, & Fukuyama, 2004; Osaka & Osaka, 2005, 2009) which show that Japanese ideophones activate the relevant sensory cortical areas. Ideophones expressing laughter activate the “laughter module” (Osaka et al., 2003) across the visual cortex, extrastriate cortex, and the premotor cortex, and also the striatal reward area. Ideophones expressing pain (e.g. *chikuchiku* for a needle-prick kind of pain, *gangan* for a throbbing headache) activate the cingulate cortex, the part of the brain which also processes actual pain. Ideophones expressing crying (e.g. oioi for *wailing*, *mesomeso* for snivelling) activate similar areas to the laughter ideophones, suggesting that crying and laughing are processed as positive and negative equivalents, but they also activate the inferior frontal gyrus and anterior cingulate cortex in the same way as the pain ideophones, suggesting that implied crying “involves some degree of concomitant emotional pain” (Osaka, 2011). Ideophones suggestive of gaze direction and manner of walking activate the frontal eye field and extrastriate visual cortex respectively. All of these ideophones activate the visual cortex and premotor cortex, which Osaka et al. argue is responsible for the vividness of the mental imagery conjured up by ideophones. However, the main limitation with these studies is that they all compared ideophones to non-words. As arbitrary words will also activate relevant sensory areas of the cortex when compared with non-words (Zwaan, 2004), this is uninformative about the special properties of sound symbolism.

Ideophones versus arbitrary words in natural language

Two neuroimaging studies have directly compared ideophones and arbitrary words. Lockwood and Tuomainen (2015) used EEG to investigate the difference between ideophonic adverbs and arbitrary adverbs by presenting Japanese speakers with sentences where the only difference was whether the adverb was sound-symbolic or not. Participants performed an unrelated sentence judgement task and were unaware of the nature of the experiment. The ideophones elicited a greater P2 and a late positive complex, both of which are in line with Arata et al.'s (2010) and Asano et al.'s (2013, 2015) findings. Lockwood and Tuomainen argue that the greater P2 in response to the ideophones represents the multisensory integration of sound and sensory processing. They also claim that while this effect is due to cross-modal associations rather than representative of true synaesthesia, the same neural mechanisms may be involved. They speculate that it is the distinctive phonological profile of ideophones which enables, or engages, the multisensory integration process. This is also in line with the conclusions of Occelli et al.'s (2013) behavioural study on autistic participants.

Kanero et al. (2014) performed two fMRI studies where participants watched animations while simultaneously hearing ideophones or arbitrary words with related to a particular modality — motion in the first experiment and shape in the second. They observed that words which participants rated as closely matching the animations elicited greater activation across the cortex than low-match words. The right posterior superior temporal sulcus (rpSTS) was activated specifically in response to ideophone trials, and not arbitrary word trials. Kanero et al. take this to mean that the right posterior STS is a critical hub for processing Japanese ideophones, and possibly sound symbolism in general. They argue that this goes beyond simple embodiment, as the rpSTS is not a perceptual or sensorimotor area related to the word meaning. Instead, Kanero et al. suggest that ideophones have a dual nature; part arbitrary linguistic symbol, part iconic symbol, and that the posterior STS works as a hub of multimodal integration. This is in line with a long tradition in the ideophone literature that emphasises the combination of iconic aspects (such as vowel size contrasts) and arbitrary aspects (such as conventional word forms) in ideophones (e.g., Diffloth, 1994). However, as ideophones contain both sensory and conventional sound symbolism, it is difficult to tease apart the separate contributions of each type with native speakers.

There has also been a study which used fMRI and fractional anisotropy (FA) to investigate sound symbolism in apparently arbitrary words. Using the same antonym stimuli and experimental set-up as Namy et al. (submitted for publication) and Tzeng et al. (submitted for publication), Revill et al. (2014) found that there was increased activation in the left superior parietal cortex in response to words which participants found sound-symbolic compared to words which they did not. Furthermore, they found a correlation between functional anisotropy in the left superior longitudinal fasciculus and participants' individual sensitivity to sound symbolism. Revill et al. argue that sound-symbolic words engage cross-modal sensory integration to a greater extent than arbitrary words, and that this cross-modal sensory integration is what facilitates word to meaning mappings (although due to the caveats mentioned above, it is not quite clear what kind of sound symbolism is under investigation here). They also argue that these correspondences may reflect some form of iconicity or embodiment, but do not speculate whether the main driver of the sound-symbolic effect is acoustic or articulatory.

Finally, Meteyard et al. (2015) investigated the phonological and semantic basis of iconicity with aphasic patients, and used it to addressed theoretical questions rather than just demonstrating an effect. They tested left-hemisphere aphasic patients with four aphasia assessment tests which assess phonology, semantics, and the combination of phonology and semantics, and looked at the processing differences between iconic and non-iconic English words (which are mostly conventionally sound-symbolic with some sensory sound-symbolic properties). Aphasics had an especially consistent processing advantage for iconic words in auditory lexical decision and reading aloud tasks, which specifically involve the mapping between phonology and semantics rather than either phonology or semantics alone. They present two potential theoretical implications, which are not mutually exclusive. Firstly, iconic words may have additional connections from the semantic system to modality-specific features, meaning that iconic words are more robust in aphasic patients because they are represented with greater redundancy within the language system itself. This means that the iconic word processing advantage is protected from damage in a similar way to high frequency, high imageability, and early acquired words. Alternatively, iconic words may be represented by direct connections between phonological form and modality-specific information. This is in line with both the embodiment semantics literature, which claims that iconic words have an extra route to activate experience, and the neuroimaging work of Kanero et al. (2014); under this account, the iconic word processing advantage in aphasics is because iconic words are additionally processed in cross-modal integration brain areas, including right hemisphere regions which are unaffected by left hemisphere damage. This study is probably the best account of how iconicity mediates between semantics and phonology rather than being specific to one or both.

Summary and future directions

The wealth of research on sound symbolism in the last few years has consolidated three main findings. Firstly, people consistently make multiple cross-modal sensory associations to specific sounds under experimental conditions, and the direction of the cross-modal sensory association — light or dark, fast or slow, etc. — is related to vowel height, vowel size, and consonant voicing of the sounds involved. Secondly, people can consistently guess the meanings of sound-symbolic words in foreign languages at an above chance level, and that this is related to phonemes and prosody. Thirdly, children are sensitive to sound symbolism and that ideophones help children acquire verbs (or at least, verbal meanings in the domain of motion) regardless of which language they are learning, meaning that children's sensitivity to ideophones is likely to be driven by the sound-symbolic phonemes and prosody. There are not yet enough neuroimaging experiments on sound symbolism to make solid conclusions, but so far it appears that sound-symbolic words activate sensory areas more strongly than arbitrary words and that the processing of sound-symbolic words appear to involve some kind of multisensory integration (or at least more multisensory integration when compared to arbitrary words).

From observation to explanation to mechanisms

The vast majority of these studies have focused on showing that there *is* an effect and have strongly made the case for sound symbolism; the next step is to investigate how this effect *works*. Prior work has supplied several important pieces of the puzzle. There are linguistic typologies and frameworks for understanding sound symbolism, such as those of Hinton et al. (2006), Perniss and Vigliocco (2014), Dingemanse (2012), and Cuskley and Kirby (2013). There are some cognitive accounts of structure mapping (Gentner, 1983), of the mental faculties for sound symbolism (Marks, 1978; Ramachandran & Hubbard, 2001), and of how sound symbolism scaffolds language acquisition (Imai & Kita, 2014). There is also a host of psychological evidence from cross-modal correspondences. However, two crucial missing pieces in the literature are specific hypotheses of how neural mechanisms may support sound symbolism, and solid neuroimaging evidence which tests them.

Broadly speaking, psychological studies have addressed the question of which particular sounds have which particular cross-modal correspondences, while linguistic studies have addressed the question of what properties sound symbolic words have which make them sound-symbolic. The current challenge in sound symbolism research is to pull together the different strands of research into one coherent field. Linguistic, psychological, and cognitive research programmes have individually made predictions about the form, use, and function of sound symbolism; this is now a perfect opportunity for cross-disciplinary collaboration to develop a neuroscientific model of sound symbolism which makes predictions that can be empirically tested with neuroimaging methods.

Interdisciplinary integration

One attempt at interdisciplinary integration is when Ramachandran and Hubbard (2001) used the *bouba/kiki* paradigm to inform their more general synaesthetic bootstrapping model of language evolution. They postulate that there is a synaesthetic correspondence between visual object shape represented in the inferior temporal lobe and sound represented in the auditory cortex, and that this synaesthetic correspondence may either happen through direct cross-activation or may be mediated by the angular gyrus. The first possibility has been interpreted as predicting that relevant sensory areas would be more strongly activated for sound-symbolic words compared to arbitrary words; the second possibility predicts that the angular gyrus would be more strongly activated for sound-symbolic words compared to arbitrary words. Both of these hypotheses can be built on with further neuroimaging work, but of the sound symbolism experiments that do mention it, they tend either show that there is a significant effect and move on, or they hedge their conclusions by suggesting that there may be a synaesthetic or embodiment mechanism without elaborating on how it might work.

Perniss and Vigliocco (2014) also provide a relatively fleshed out model. They propose that iconicity exists to provide the link between linguistic form and human experience by establishing reference and displacement through sensori-motor embodiment of linguistic form, and that the cross-linguistic variability in iconicity shows how different languages strike a balance between two basic constraints — the need to link language to human experience and the need for an efficient communication system. This suggestion provides fertile ground for hypothesis testing, especially with language development literature which can be framed in terms of investigating the emergence of reference and displacement with respect to iconicity. The next step for this model is to hypothesise how the brain processes sound symbolism and cross-modal correspondences. Perhaps there is a role here Meteyard et al. (2015)'s suggestion that iconic words may be supported by additional connectivity between semantic or phonological representations and perceptuo-motor information.

Recent research on sound symbolism has established that sound symbolism is widespread across languages, that it has cross-modal correspondences with other senses, that this has an effect on behaviour and development, and that it elicits distinct brain signals. We are now at an exciting juncture where we can start approaching this phenomenon from an integrated interdisciplinary perspective. Ideophones and sound symbolism from natural languages provide us with opportunities for the experimental investigation of the role of sound symbolism in meaning, interpretation, and perception. Through hypothesis-testing and model-building, these experiments will help contribute to a cumulative science of sound symbolism in human language.

References

Akita, K. (2009a). A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics. Kobe University.

Akita, K. (2009b). The acquisition of the constraints on mimetic verbs in Japanese and Korean. In Y. Takubo (Ed.), Japanese/Korean linguistics (Vol. 16, pp. 163–177). Stanford, CA: CSLI Publications.

Arata, M., Imai, M., Kita, S., Thierry, G., & Okada, H. (2010). Perception of sound symbolism in 12 month-old infants: An ERP study. Neuroscience Research, 68(Abstracts of the 33rd Annual Meeting of the Japan Neuroscience Society (Neuro 2010)), e300.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. Cortex, 63, 196–205. http://doi.org/10.1016/j.cortex.2014.08.025

Asano, M., Kitajo, K., Thierry, G., Kita, S., Okada, H., & Imai, M. (2013). Linguistic experience alters the processing of sound symbolism: An ERP study. Presented at the Sound symbolism Workshop 2013, Tokyo.

Asano, M., & Yokosawa, K. (2011). Synesthetic colors are elicited by sound quality in Japanese synesthetes. Consciousness and Cognition, 20(4), 1816–1823. http://doi.org/10.1016/j.concog.2011.05.012

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. Memory & Cognition, 40(1), 83–92. http://doi.org/10.3758/s13421-011-0139-3

Bankieris, K., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? Cognition, 136, 186–195. http://doi.org/10.1016/j.cognition.2014.11.013

Bergen, B. K. (2004). The Psychological Reality of Phonaesthemes. Language, 80(2), 290–311. http://doi.org/10.2307/4489664

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). 'Bouba' and 'Kiki' in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. Cognition, 126(2), 165–172. http://doi.org/10.1016/j.cognition.2012.09.007

Bühler, K. (1990). Theory of Language: The Representational Function of Language. Amsterdam: John Benjamins Publishing.

Corballis, M. C. (2002). From hand to mouth: the origins of language. Princeton, N.J.; Woodstock: Princeton University Press.

Cuskley, C. (2013). Mappings between linguistic sound and motion. The Public Journal of Semiotics, 5(1), 39–62.

Cuskley, C., & Kirby, S. (2013). Synaesthesia, cross-modality, and language evolution. In J. Simner & E. M. Hubbard (Eds.), Oxford Handbook of Synaesthesia (pp. 869–907). Oxford: Oxford University Press.

Davis, R. (1961). The Fitness of Names to Drawings. a Cross-Cultural Study in Tanganyika. British Journal of Psychology, 52(3), 259–268. http://doi.org/10.1111/j.2044-8295.1961.tb00788.x

Diffloth, G. (1994). i: big, a: small. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), Sound Symbolism (pp. 107–114). Cambridge: Cambridge University Press.

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. Language and Linguistics Compass, 6(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M. (2013). Ideophones and gesture in everyday speech. Gesture, 13(2), 143–165. http://doi.org/10.1075/gest.13.2.02din

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What Sound Symbolism Can and Cannot Do: Testing the Iconicity of Ideophones from Five Languages. Language, 92(2), e117-33. http://doi.org/10.1353/lan.2016.0034

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound symbolism is disrupted in dyslexia: implications for the role of cross-modal abstraction processes. In Proceedings of the 37th Annual Meeting of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Elsen, H. (1991). Erstspracherwerb, Der Erwerb des deutschen Lautsystems. Wiesbaden, Germany: Deutscher Universitäts Verlag.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. Proceedings of the National Academy of Sciences, 103(32), 12203–12208. http://doi.org/10.1073/pnas.0602173103

Fernández-Prieto, I., Navarra, J., & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. Infant Behavior and Development, 38, 77–81. http://doi.org/10.1016/j.infbeh.2014.12.008

Fort, M., Weiss, A., Martin, A., & Peperkamp, S. (2013). Looking for the bouba-kiki effect in prelexical infants (pp. 71–76). Presented at the 12th International Conference on Auditory-Visual Speech Processing, Annecy.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. Trends in Cognitive Sciences, 6(2), 78–84. http://doi.org/10.1016/S1364-6613(00)01839-8

Gasser, M., Sethuraman, N., & Hockema, S. (2010). Iconicity in expressives: an empirical investigation. In S. Rice & J. Newman (Eds.), Experimental and Empirical Methods in the Study of Conceptual Structure, Discourse, and Language (pp. 163–180). Stanford, CA: CSLI Publications.

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy*. Cognitive Science, 7(2), 155–170. http://doi.org/10.1207/s15516709cog0702_3

Gentner, D., & Boroditsky, L. (2001). Individuation, Relativity, and Early Word Learning. In M. Bowerman & S. C. Levinson (Eds.), Language Acquisition and Conceptual Development (pp. 215–256). Cambridge: Cambridge University Press.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. Frontiers in Language Sciences, 4, 416. http://doi.org/10.3389/fpsyg.2013.00416

Hamano, S. (1998). The Sound-Symbolic System of Japanese. Center for the Study of Language and Information ; Tokyo.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33(2–3), 61–83. http://doi.org/10.1017/S0140525X0999152X

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. Nature Reviews Neuroscience, 8(5), 393–402. http://doi.org/10.1038/nrn2113

Hinton, L., Nichols, J., & Ohala, J. J. (1994). Sound Symbolism. Cambridge: Cambridge University Press.

Hinton, L., Nichols, J., & Ohala, J. J. (2006). Sound Symbolism. Cambridge: Cambridge University Press.

Hirata, S., Ukita, J., & Kita, S. (2011). Implicit phonetic symbolism in voicing of consonants and visual lightness using Garner's speeded classification task. Perceptual and Motor Skills, 113(3), 929–940. http://doi.org/10.2466/15.21.28.PMS.113.6.929-940

Hockett, C. (1959). Animal 'languages' and human language. Human Biology, 31(1), 32–39.

Hockett, C. (1960). The Origin of Speech. Scientific American, 203, 89–97.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130298. http://doi.org/10.1098/rstb.2013.0298

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. Cognition, 109(1), 54–65. http://doi.org/10.1016/j.cognition.2008.07.015

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007a). How does it hurt, kiri-kiri or siku-siku? Japanese mimetic words of pain perceived by Japanese speakers and English speakers. In M. Minami (Ed.), Applying theory and research to learning Japanese as a foreign language (pp. 2–19). Newcastle: Cambridge Scholars Publishing. Retrieved from http://eprints.soas.ac.uk/6438/

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007b). What do English speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. Japanese-Language Education around the Globe, 17, 53–78.

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How Sound Symbolism Is Processed in the Brain: A Study on Japanese Mimetic Words. PLoS ONE, 9(5), e97905. http://doi.org/10.1371/journal.pone.0097905

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese Sound symbolism Facilitates Word Learning in English-Speaking Children. Cognitive Science, 35(3), 575–586. http://doi.org/10.1111/j.1551-6709.2010.01169.x

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. Linguistics, 35(2), 379–415. http://doi.org/10.1515/ling.1997.35.2.379

Klamer, M. (2002). Semantically Motivated Lexical Patterns: A Study of Dutch and Kambera Expressives. Language, 78(2), 258–286.

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

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. Cognition, 114(1), 19–28. http://doi.org/10.1016/j.cognition.2009.08.016

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. Journal of Verbal Learning and Verbal Behavior, 10(4), 427–429. http://doi.org/10.1016/S0022-5371(71)80042-7

Kwon, N., & Round, E. R. (2014). Phonaesthemes in morphological theory. Morphology, 25(1), 1–27. http://doi.org/10.1007/s11525-014-9250-z

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. First Language, 34(5), 387–405. http://doi.org/10.1177/0142723714550110

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. Behavioral and Brain Sciences, 22(1), 1–38.

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. Language Sciences, 6, 933. http://doi.org/10.3389/fpsyg.2015.00933

Maglio, S. J., Rabaglia, C. D., Feder, M. A., Krehm, M., & Trope, Y. (2014). Vowel Sounds in Words Affect Mental Construal and Shift Preferences for Targets. Journal of Experimental Psychology: General, 143(3), 1082–1096. http://doi.org/10.1037/a0035543

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., … Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. Cognition, 114(3), 299–319. http://doi.org/10.1016/j.cognition.2009.10.002

Majid, A. (2012). Current Emotion Research in the Language Sciences. Emotion Review, 4(4), 432–443. http://doi.org/10.1177/1754073912445827

Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. Cognition, 130(2), 266–270. http://doi.org/10.1016/j.cognition.2013.11.004

Malkiel, Y. (1994). Regular Sound Development, Phonosymbolic Orchestration, Disambiguation of Homonyms. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), Sound Symbolism (pp. 207–21). Cambridge: Cambridge University Press.

Marks, L. E. (1978). The unity of the senses: interrelations among the modalities. New York; London: Academic Press.

Maurer, D., & Mondloch, C. J. (2004). Neonatal synesthesia: A re-evaluation. In L. Robertson & N. Sagiv (Eds.), Attention on Synesthesia: Cognition, Development and Neuroscience (pp. 193–213). Oxford: Oxford University Press.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound–shape correspondences in toddlers and adults. Developmental Science, 9(3), 316–322. http://doi.org/10.1111/j.1467-7687.2006.00495.x

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. Neuropsychologia. http://doi.org/10.1016/j.neuropsychologia.2015.01.042

Monaghan, P., & Christiansen, M. H. (2006). Why form-meaning mappings are not entirely arbitrary in language. In Proceedings of the 28th annual conference of the Cognitive Science Society (pp. 1838–1843). Mahwah, NJ: Lawrence Erlbaum.

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. Journal of Experimental Psychology: General, 140(3), 325–347. http://doi.org/10.1037/a0022924

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130299. http://doi.org/10.1098/rstb.2013.0299

Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: Vowel colours in the ear of the beholder. I-Perception, 5(2), 132–142. http://doi.org/10.1068/i0626

Namy, L. L., DeFife, L. C., Mathur, N. M., & Nygaard, L. C. (submitted for publication). Cross-linguistic sound symbolism: Phonetic determinants of word meaning. Submitted for Publication.

Newman, S. S. (1933). Further Experiments in Phonetic Symbolism. The American Journal of Psychology, 45(1), 53. http://doi.org/10.2307/1414186

Newmeyer, F. J. (1992). Iconicity and Generative Grammar. Language, 68(4), 756. http://doi.org/10.2307/416852

Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 65(2), 115–124. http://doi.org/10.1037/a0022268

Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. Language and Cognition, 4(2), 115–125. http://doi.org/10.1515/langcog-2012-0007

Nielsen, A., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 67(2), 153–163. http://doi.org/10.1037/a0030553

Nuckolls, J. B. (1999). The Case for Sound Symbolism. Annual Review of Anthropology, 28(1), 225–252. http://doi.org/10.1146/annurev.anthro.28.1.225

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. Cognition, 112(1), 181–186. http://doi.org/10.1016/j.cognition.2009.04.001

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. Cognitive Science, 33(1), 127–146. http://doi.org/10.1111/j.1551-6709.2008.01007.x

Occelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013). The takete–maluma phenomenon in autism spectrum disorders. Perception, 42(2), 233 – 241. http://doi.org/10.1068/p7357

Oda, H. (2000). An embodied semantic mechanism for mimetic words in Japanese. Indiana University, Indiana.

Ohtake, Y., & Haryu, E. (2013). Investigation of the process underpinning vowel-size correspondence. Japanese Psychological Research, 55(4), 390–399. http://doi.org/10.1111/jpr.12029

Osaka, N. (2009). Walk-related mimic word activates the extrastriate visual cortex in the human brain: an fMRI study. Behavioural Brain Research, 198(1), 186–189. http://doi.org/10.1016/j.bbr.2008.10.042

Osaka, N. (2011). Ideomotor response and the neural representation of implied crying in the human brain: An fMRI study using onomatopoeia1. Japanese Psychological Research, 53(4), 372–378. http://doi.org/10.1111/j.1468-5884.2011.00489.x

Osaka, N., & Osaka, M. (2005). Striatal reward areas activated by implicit laughter induced by mimic words in humans: a functional magnetic resonance imaging study. Neuroreport, 16(15), 1621–1624.

Osaka, N., & Osaka, M. (2009). Gaze-related mimic word activates the frontal eye field and related network in the human brain: an fMRI study. Neuroscience Letters, 461(2), 65–68. http://doi.org/10.1016/j.neulet.2009.06.023

Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., & Shibasaki, H. (2003). An emotion-based facial expression word activates laughter module in the human brain: a functional magnetic resonance imaging study. Neuroscience Letters, 340(2), 127–130. http://doi.org/10.1016/S0304-3940(03)00093-4

Osaka, N., Osaka, M., Morishita, M., Kondo, H., & Fukuyama, H. (2004). A word expressing affective pain activates the anterior cingulate cortex in the human brain: an fMRI study. Behavioural Brain Research, 153(1), 123–127. http://doi.org/10.1016/j.bbr.2003.11.013

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. Journal of Experimental Child Psychology, 114(2), 173–186. http://doi.org/10.1016/j.jecp.2012.05.004

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. Frontiers in Psychology, 1(E227). http://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130300. http://doi.org/10.1098/rstb.2013.0300

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia -- A window into perception, thought and language. Journal of Consciousness Studies, 8(12), 3–34.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. Brain and Language, 128(1), 18–24. http://doi.org/10.1016/j.bandl.2013.11.002

Sapir, E. (1929). A study in phonetic symbolism. Journal of Experimental Psychology, 12(3), 225–239. http://doi.org/10.1037/h0070931

Saussure, F. de. (1916). Course in general linguistics (1959 edition). New York: Philosophical Library.

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. Language Sciences, 5, 80. http://doi.org/10.3389/fpsyg.2014.00080

Sedley, D. N. (2003). Plato’s Cratylus. Cambridge: Cambridge University Press.

Shin, E., & Kim, C.-Y. (2014). Both ‘나’ and ‘な’ are yellow: Cross-linguistic investigation in search of the determinants of synesthetic color. Neuropsychologia. http://doi.org/10.1016/j.neuropsychologia.2014.09.032

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. Perception, 39(4), 553–569.

Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. I-Perception, 4(4), 239–241. http://doi.org/10.1068/i0535

Thompson, P. (2013, February). Phonetic symbolism for size, shape, and motion (PhD thesis). University of Warwick, Warwick. Retrieved from http://webcat.warwick.ac.uk/record=b2715869~S1

Thompson, P., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. The Quarterly Journal of Experimental Psychology, 64(12), 2392–2404. http://doi.org/10.1080/17470218.2011.605898

Tzeng, C., Nygaard, L. C., & Namy, L. L. (submitted for publication). The specificity of sound symbolic correspondences in spoken language. Submitted for Publication.

Voeltz, F. K. E., & Kilian-Hatz, C. (2001). Ideophones. Amsterdam: John Benjamins Publishing.

von Humboldt, W. (1836). On Language: On the Diversity of Human Language Construction and Its Influence on the Mental Development of the Human Species. (M. Losonsky, Ed.). Cambridge: Cambridge University Press.

Waugh, L. R. (1994). Degrees of iconicity in the lexicon. Journal of Pragmatics, 22(1), 55–70. http://doi.org/10.1016/0378-2166(94)90056-6

Yoshida, H. (2012). A Cross-Linguistic Study of Sound Symbolism in Children’s Verb Learning. Journal of Cognition and Development, 13(2), 232–265. http://doi.org/10.1080/15248372.2011.573515

Zwaan, R. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In Psychology of learning and motivation (Vol. 44, pp. 35–62).

Chapter 3

Ideophones in Japanese modulate the P2 and late positive complex responses.

Published as:
Lockwood & Tuomainen (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. Frontiers in Psychology 6.

Abstract

Sound-symbolism, or the direct link between sound and meaning, is typologically and behaviourally attested across languages. However, neuroimaging research has mostly focused on artificial pseudowords or individual segments, which do not represent sound-symbolism in natural language. We used EEG to compare Japanese ideophones, which are phonologically distinctive sound-symbolic lexical words, and arbitrary adverbs during a sentence reading task. Ideophones elicit a larger visual P2 response than arbitrary adverbs, as well as a sustained late positive complex. Our results and previous literature suggest that the larger P2 may indicate the integration of sound and sensory information by association in response to the distinctive phonology of ideophones. The late positive complex may reflect the facilitated lexical retrieval of arbitrary words in comparison to ideophones. This account provides new evidence that ideophones exhibit similar cross-modal correspondences to those which have been proposed for pseudowords and individual sounds.

Introduction

Sound-symbolism, most simply defined as “the direct linkage between sound and meaning” (Hinton, Nichols, & Ohala, 2006: p.1), has traditionally played a peripheral role in linguistics. The assumption of an iconic link between form and meaning conflicts profoundly with the principle of arbitrariness (Saussure, 1959), which holds language as “a wholly symbolic system, [whereby] the elements of which are manipulated on an abstract level of representation” (Perniss, Thompson, & Vigliocco, 2010: p.1). The progression of formal linguistics is a fair indication that its foundations of arbitrariness are solid, but this is not to say that iconic and arbitrary systems in language cannot exist side by side. Wider recognition of the existence and extent of such words has been hindered until relatively recently by two main factors; a historically Eurocentric linguistic perspective, which has contributed to the assumption that the relative paucity of sound-symbolism in most Indo-European languages is reflective of human language as a whole (Perniss et al., 2010), and a disunity of description and definition (Dingemanse, 2012). However, the field has a rather more global perspective than it used to, and has settled on *ideophone* as the preferred term for lexical classes of words which exhibit non-arbitrary relations between their form and their meaning. Ideophones are more specifically defined as "marked words that depict sensory imagery" (Dingemanse, 2012), and are found in various natural languages (Childs, 1994; Diffloth, 1972; Dingemanse, 2012; Firth, 1964; Hamano, 1998; Perniss et al., 2010).

The study of sound-symbolism is enjoying a renaissance in psychology too. The best known example of purely sensory sound-symbolism is Ramachandran and Hubbard’s (2001) *bouba/kiki* experiment, based on Köhler’s (1947) observations using *maluma/takete*. Participants overwhelmingly linked the novel word *kiki* with a spiky shape and *bouba* with a round shape, regardless of their native language. Ramachandran and Hubbard argued that the connections which humans make between voiceless consonants and sharp contours, and between voiced consonants and round contours (as well as between vowels and object size (Johnson, 1967; Ohala, 1994)), are synaesthetic. Many similar behavioural studies followed, which focused on eliciting participants’ impressions of nonsense words (Maurer et al., 2006; Ramachandran & Hubbard, 2001), asking participants to associate tastes to vowel sounds (Simner, Cuskley, & Kirby, 2010), or investigating the names of cancer drugs and patients’ perceptions of their effectiveness (Abel & Glinert, 2008). Such behavioural studies are certainly interesting, and have shown that form to meaning mappings cover various different sensory systems and are relatively consistent across languages, which means that speakers of one language to apply the same correspondences to other languages with some degree of success (Revill, Namy, DeFife, & Nygaard, 2014). However, apart from a few studies on the perception of ideophones by naïve speakers (Berlin, 1994; Iwasaki, Vinson, & Vigliocco, 2007b; Oda, 2000), the extensive ideophone literature in linguistics has mostly been overlooked. Very few studies use naturally-occurring sound-symbolic words, and instead rely on nonsense words which have been deliberately constructed to maximise perceptual differences between conditions. While informative about the nature of synaesthetic effects between sounds and various sensory systems, this does not reflect sound-symbolism as it is used in real language.

Research in Japanese linguistics has proven to be more fruitful for examining sound-symbolism using natural language. Japanese has an extensive set of ideophones (known as *mimetics* in the Japanese literature) (Akita, 2009; Hamano, 1998; Kita, 1997), which lend themselves to behavioural, developmental, and neuroimaging experiments where more specific hypothesis can be tested. Several studies have shown that naïve participants are sensitive to the association between ideophone form and ideophone meaning at above chance level, and that articulation facilitates this connection between form and meaning (Iwasaki et al., 2007b; Kunihira, 1971; Oda, 2000). However, Japanese sound-symbolism is not entirely transparent to other speakers. Iwasaki et al. (2007b) found that English speakers with no knowledge of Japanese could make accurate judgments about ideophones expressing certain semantic dimensions but not others (a finding which has been supported by corpus studies, such as Gasser et al. (2010), which have shown a correlation between phonetic form and semantic domain in ideophones). This suggests that, while there is some solid evidence for cross-linguistic sound-symbolism, ideophones are not completely intuitive to speakers of other languages. Developmental literature suggests that this general cross-linguistic insensitivity to ideophones may be conditioned in adults by language’s arbitrary and conventionalised connections. Children are more sensitive to sound-symbolic patterns in Japanese than to arbitrary forms, and this sensitivity appears to function as a scaffold for language acquisition (Ishiguro, 1993; Imai, Kita, Nagumo, & Okada, 2008; Iwasaki et al., 2007b). This finding is consistent across both Japanese and English-speaking children (Kantartzis, Imai, & Kita, 2011; Yoshida, 2012), providing evidence towards a cross-linguistic – or, perhaps more accurately, language-independent – early sensitivity towards sound-symbolism.

EEG studies on sound-symbolism have found a variety of different effects at a variety of different time points, and so it is still unclear what the general effects of sound-symbolism are. Most EEG studies have used matching tasks, where participants are presented with images and sound-symbolic pseudowords which are either sound-symbolically congruent (e.g. *kiki* and a spiky shape) or incongruent. Kovic, Plunkett, and Westermann (2010) found that congruent conditions elicited a greater negative-going wave at 140-180ms, and that this effect was most prominent at occipital electrodes. Asano et al. (2015) tested preverbal 11-month-old infants, and found that incongruent conditions elicited greater N400 responses. Moreover, phase synchronization of neural oscillations increased more in incongruent conditions around 400ms, suggesting that incongruent conditions required sustained effort for cross-modal binding. They also found amplitude increases in the gamma band in centro-parietal within 300ms of word onset for congruent conditions, suggesting that 11-month-olds process sound-symbolism as perceptual binding. Bien et al.'s (2012) study on cross-modal binding appears to be consistent with Asano et al.'s findings. Bien et al. investigated pitch-size mappings rather than linguistic sound-size mappings, and found that cross-modally congruent conditions (i.e. high pitch and small size) elicited a greater P2 response at around 250ms in intra-parietal regions. Bien et al. argue that this effect is synaesthetic, and that cross-modal mappings underlie and influence multisensory perception. Taken together, the Asano et al. (2015) and Bien et al. (2012) studies suggest that cross-modally congruent conditions show an early effect in centro-parietal areas which is related to sensory processing, while incongruent conditions show a later effect which is related to semantic integration difficulty. The Kovic et al. (2010) study found a still earlier effect for congruent conditions, but at occipital regions, and moreover was a negative-going rather than positive-going wave. This may be related to the fact that Kovic et al. (2010) trained their participants to learn the mappings first, rather than requiring spontaneous decisions.

The P2 may well be affected by sound-symbolism. In general, the P2 response has been linked to multisensory integration (Bien et al., 2012). In single modality experiments, the auditory P2 response has been linked to higher level categorization processes (Crowley & Colrain, 2004), and auditory processing of linguistic (Carpenter & Shahin, 2013) and non-linguistic information (Shahin, Roberts, Pantev, Trainor, & Ross, 2005). The visual P2 elicited in word reading studies has been linked to phonological and semantic analysis (Dien, 2009; Kong et al., 2010; Landi & Perfetti, 2007). However, the P2 remains relatively poorly understood and is likely to represent several underlying component generation processes, so any functional significance claims regarding the P2 should be taken with caution. Moreover, the topographic distribution of the P2 is not well-established; some studies find a visual P2 elicited in parieto-occipital areas (Freunberger, Klimesch, Doppelmayr, & Höller, 2007), while others find a more frontal distribution (Barber, Ben-Zvi, Bentin, & Kutas, 2011; Key, Dove, & Maguire, 2005). As sound-symbolism is related to the expression of other sensory information and may well recruit these senses during language processing, ERP investigations into multisensory processing are also important to consider. Many of these have centred around visual object naming, recognition, or categorization with respect to congruent and incongruent auditory stimuli (Bien et al., 2012; Giard & Peronnet, 1999; Molholm et al., 2002; Molholm, Ritter, Javitt, & Foxe, 2004; Widmann, Gruber, Kujala, Tervaniemi, & Schröger, 2007; Widmann, Kujala, Tervaniemi, Kujala, & Schröger, 2004; Yuval-Greenberg & Deouell, 2007), but others have hinted at its role in synaesthetic experiences (Brang, Kanai, Ramachandran, & Coulson, 2010). Neuroimaging approaches towards ideophones provide evidence for a strong synaesthetic effect between ideophone meaning and brain activation in relevant, non-linguistic areas, while showing that pseudowords do not elicit similar activity (Osaka & Osaka, 2005, 2009; Osaka, 2009, 2011; Osaka et al., 2003; Osaka et al., 2004). These studies argue that the vividness of the mental imagery conjured up by ideophones is due to this ideomotor response across the visual and premotor cortices. These findings suggest a neurological basis for the early definitions (such as Doke 1935) of ideophones as "vivid". Kanero et al. (2014) used locomotion and shape ideophones in contrast to arbitrary adverbs and verbs. They found that ideophones activated the right posterior STS, and argue that this area may be the primary location for processing sound-symbolism. They argue that this may also reflect how sound symbolic words function as both linguistic and non-linguistic iconic symbols, while also providing support for hypotheses that sound-symbolism involves cross-domain mapping (Davis, 1961; Martino & Marks, 2001; Ramachandran & Hubbard, 2001).

This study uses naturally-occurring ideophones rather than deliberately-constructed pseudowords to investigate sound-symbolism, in particular the suggestion that it is facilitated by synaesthesia-like cross modal mappings. In this study, we recorded event-related potentials (ERPs) to investigate the temporal characteristics of processing Japanese ideophones in visually-presented normal and nonsense sentences. Based on the literature reviewed above, we focused on two ERP responses, the P2 and the N400. We predicted that if significant differences at the P2 between ideophones and arbitrary adverbs were found, it would suggest that phonological processing and/or integration of sensory information in different domains may be involved in the processing of ideophones. If an N400 effect were found, it would suggest that the semantic properties of ideophones are different from those of arbitrary words. Finally, we ran a pilot study run using ten Japanese participants, which showed a P2 response and a sustained late positive response starting at around 400ms. Accordingly, the focus of the current study was on the P2, N400, and late positive complex responses.

Methods

We first carried out a pilot EEG study on a small sample of ten native Japanese speakers. The task was identical to the main experiment, and was done in order to check the stimuli and to establish whether there was any indication of an effect. Initial analyses of these ten participants confirmed that the P2 and late positive complex were areas of interest for the full experiment. Based on their feedback and corrections, the following stimuli were used for the current experiment.

Stimuli

Adverbial ideophones with a CVCV-CVCV pattern were the most suitable due to their frequency, variety of meanings, and iconicity of form. Using the Sketchengine program to access the JpWaC online corpus (Kilgarriff et al., 2004), a list of the most frequently used adverbial ideophones was created, and 35 of these were selected for use in the stimuli sentences, as recommended by Kutas et al. (2010). Full stimuli sentences of 3-5 words long were created for these 35 ideophonic adverbs (henceforth referred to as iconic adverbs), as shown in (1) below:

(1) Hanako-ha samusa-de gatagata furueta
Hanako-SUBJ cold-with shiver-ICONIC.ADV shake-PAST
‘Hanako shook with shivers because of the cold.’

Thirty-five equivalent arbitrary adverbs were then matched to the iconic adverbs as closely as possible for meaning, total word length, and frequency. Frequency ratings were obtained from the JpWaC corpus; average frequency for ideophones was 2791.9 occurrences in the corpus, average frequency for arbitrary adverbs was 3254.7 occurrences in the corpus, and this was not significantly different (t-test, t= -1.0886, p=0.284). The first character frequency in each word was also consistent across conditions (t-test, t=0.736, p=0.467), and the variety of pitch accents across the stimuli reduces any potential systematic confounds. The same sentences were used for both adverbs, resulting in two conditions - iconic and arbitrary. The iconic and arbitrary adverbs were always located after the first and before the last word in the sentence.

Two further conditions were added to create a behavioural task which would distract participants from the purpose of the experiment. Another verb from a different sentence in the stimulus set was selected in order to make a nonsense sentence in a typical N400 elicitation experiment. Participants were asked to make decisions about whether the sentence meaning was normal or strange. This resulted in 140 sentences spread across four conditions – iconic/sensible, arbitrary/sensible, iconic/nonsense, and arbitrary/nonsense, shown in Table 1 below.

Condition		Example sentence
Iconic	Sensible	Hanako-ha samusa-de gatagata furueta Hanako-SUBJ cold-with shiver-ICONIC.ADV shake-PAST ‘Hanako shook with shivers because of the cold.’
Arbitrary	Sensible	Hanako-ha samusa-de sukoburu furueta Hanako-SUBJ cold-with greatly shake-PAST ‘Hanako shook greatly because of the cold.’
Iconic	Nonsense	Hanako-ha samusa-de gatagata oshieta Hanako-SUBJ cold-with shiver-ICONIC.ADV teach-PAST ?‘Hanako shake-taught because of the cold.’
Arbitrary	Nonsense	Hanako-ha samusa-de sukoburu oshieta Hanako-SUBJ cold-with greatly teach-PAST ?‘Hanako taught greatly because of the cold.’

Table 1: experimental conditions for one set of four example stimuli sentences

The 2x2 experimental design was chosen for two reasons. Firstly, it enabled a design where the only difference between the sentences was whether or not the adverb was iconic. Secondly, it allowed us to give the participants a sensibility judgment task to keep them focused throughout the experiment. Participants were asked to judge whether the sentences were sensible or nonsense, and to respond by pressing a button accordingly.

It is important to note here that it is the verb which determines whether a sentence is a sensible or nonsense sentence. Therefore, up until the presentation of the sentence-final verb, all sentences are sensible. This means that all ideophones are presented at a point during the sentence where it makes sense, and so the only condition contrast of interest when analysing the effects of sound-symbolism is whether the word is iconic or arbitrary.

A further 140 filler sentences with no adverbs were also included, giving 280 trials in total. The filler sentences were divided between intransitive, transitive, and ditransitive sentences, and were also split into sensible and nonsense conditions. This further served to disguise the underlying purpose of the experiment. These filler sentences were mostly sourced from dictionary examples.

Participants

The experiment was carried out on 22 adult Japanese native speaking participants (18f, 4m) aged 19 to 31 with normal or corrected to normal vision. 20 were right-handed, two were left-handed. Participants were recruited from various London universities and were reimbursed with £15 for their time. All participants were educated within the Japanese school system, and are therefore highly proficient readers of Japanese. Ethics approval was obtained from the UCL Research Department of Linguistics Ethics Committee (project ref LING-2013-06-25).

Procedure

The experiment was conducted in a sound-attenuated, dimly-lit room. Participants sat in a chair facing a 17-inch screen situated approximately 90cm away. Before each session, participants were given a short practice block. The sentences in the experimental block were randomised and divided into seven blocks of 40 sentences. The order of these blocks was randomised for each participant. Each block lasted approximately four minutes, and participants were given a short break between each block.

The sentences were presented visually word by word in the centre of the screen. Each sentence was preceded by a fixation cross, whose duration was randomised to be between 1900ms and 2100ms. The jitter was included in order to reduce contingent negative variation, a low frequency negative wave elicited when participants expect a stimulus (Luck, 2005). Each word was presented for 1000ms, and a blank screen was presented between each word for 100ms. Participants were asked to read the whole sentence, and then decide by a button press whether the sentence was sensible or nonsense after the last word was presented. There were 140 sentences which required “sensible” responses and 140 which required “nonsense” responses. Participants were asked to blink and move between trials but to stay as still as possible during trials.

Words were generally presented in the most natural Japanese script; that is, Jōyō kanji were used for verbs, adjectives, and nouns, katakana for loanwords, and hiragana for adverbs, ideophones, and grammatical functions. Written sentences were double checked with a native speaker, and then refined with ten more native speakers in the pilot test. There were some disagreements between native speakers over the naturalness of some sentences, but all sentences were considered acceptable before the presentation of the final verb. We therefore do not expect to find a main effect of sentence sense in ERPs taken at the presentation of the ideophone or arbitrary adverb.

EEG Recording and Analysis

The continuous EEG was recorded using 64 electrodes fixed to an elastic cap (Biosemi; (http://www.biosemi.com/headcap.htm)) using the 10-10 system. All electrode offsets were kept within ±20mV as recommended by Biosemi. All EEG and external channels were amplified using a Biosemi ActiveTwo DC amplifier. EEG-waveforms were time-locked to the onset of the presentation of the orthographic stimulus.

The electro-oculogram was recorded from two external electrodes; one below the left eye to measure blinks and vertical eye movements (VEOG), and one at the right canthus to measure horizontal eye movements (HEOG). All electrodes were referenced off-line to the average of left and right mastoids.

EEG data were filtered with 0.5-30Hz bandpass offline. An independent component analysis (ICA) implemented in EEGLAB v. 12.0.2.04b (Delorme & Makeig 2004) was run on all participants' data and horizontal and vertical eye-movements as well as blinks were removed from the data. Trials which still contained artefacts were rejected offline using the ERPLAB v. 3.0.2.1 (Lopez-Calderon & Luck, 2014) artefact detection tools. The moving window peak-to-peak threshold tool (moving window width: 200ms, voltage threshold: 100µV, window step: 20 ms) and the step-like artefacts tool (moving window width: 400ms, voltage threshold: 35 µV, window step: 10ms) were used to reject trials with these artefacts.

One participant was excluded from the analysis due to heavy sweating during the experiment making the data unusable. In the remaining 21 participants, 18.9% of critical trials were excluded. In three participants, readings for bad electrodes (F6; PO4 and PO8; POz, PO4, and PO8 in the three participants respectively) were not included in the ICA and were interpolated from adjacent electrodes before artefact detection.

Averaged ERPs were calculated per condition per participant from 200ms pre-onset to 800ms post-onset of each stimulus word. These were then grouped into critical bins for the four experimental conditions; iconic adverbs in sensible sentences, iconic adverbs in nonsense sentences, arbitrary adverbs in sensible sentences, and arbitrary adverbs in nonsense sentences.

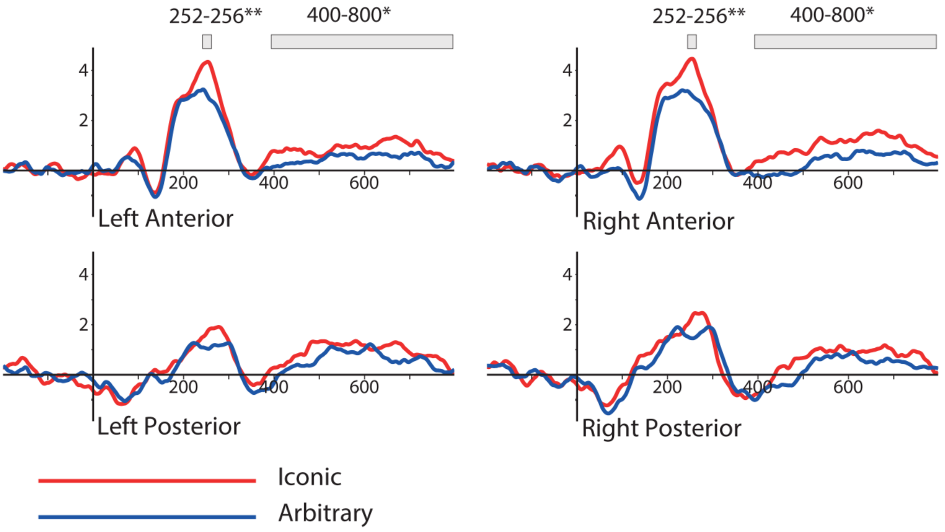
The P2 response was calculated with an automatic peak-detection procedure in the 200-300ms window from the grand average, and the resulting mean amplitudes within a 2ms interval around each peak were used for further statistical analysis. The identified peak was 254ms, and so a window of 252-256ms was used for analyses. This procedure was used in order to match the analyses in Bien et al. (2012) as closely as possible; the similar latencies of the resulting peaks (254ms and 250ms) suggest that the same response has been identified. Repeated measures 2x2x64 ANOVAs were computed to investigate three factors; iconicity, sense, and electrode location. The factor of iconicity had two levels, iconic and arbitrary. The factor of sense also had two levels, sensible and nonsense. The factor of electrode location had 64 levels, which were all electrode locations from which EEG recordings were taken. In order to get a broader picture of the distribution of the effect across the scalp, we also performed a quadrant analysis. After excluding midline electrodes, we grouped the remaining 44 electrodes into four quadrants of 11: left anterior (Fp1, AF7, AF3, F7, F5, F3, F1, FT7, FC5, FC3, FC1), right anterior (Fp2, AF4, AF8, F2, F4, F6, F8, FC2, FC4, FC6, FT8), left posterior (TP7, CP5, CP3, CP1, P7, P5, P3, P1, PO7, PO3, O1), and right posterior (CP2, CP4, CP6, TP8, P2, P4, P6, P8, PO4, PO8, O2). A repeated measures 2x2x4 ANOVA was then computed.

The N400 response was calculated more generally, taking the mean amplitude across a time window of 350-550ms. This slightly later time window was taken in order to remain consistent with existing Japanese sound-symbolic ERP literature (Asano et al., 2015). The same repeated measures 2x2x64 ANOVAs were computed to investigate three factors; iconicity, sense, and electrode location. As the data did not resemble a typical N400 response but rather a long-lasting increased positivity which appeared to begin at around 400ms and last until the end of the trial, we also tested the time window of 400-800ms in the same way.

Results

Behavioural results showed that participants made errors on 7.09% of trials. The individual error rate ranged from 2.5% to 16.8%, so no participant was excluded on the basis of high error rate. Mistakes were broadly consistent across the main conditions, with participants making mistakes on 7.41% of iconic sentences and 11.77% of arbitrary sentences. This was not significantly different (t=1.61, p=0.11). Eight sentences were miscategorised by over half the participants. One sentence was a filler sentence taken from the dictionary (Shogakukan, 2001), and seven were critical trials. However, as behavioural responses were unrelated to the underlying purpose of the experiment, and as it was the sentence-final verb which dictated the sensibility of a sentence whereas the ERP analysis was conducted on target adverbs presented before the final verbs, incorrect responses were not excluded from the iconic/arbitrary target word analysis.

In the P2 analysis, there was a significant difference between ERP amplitudes elicited by iconic and arbitrary conditions (F=10.095, df=1,20, p=0.005, partial η^2=0.335). The interaction between iconicity and electrode location only tended towards significance (F=2.045, p=0.088). In order to get a broader picture of the distribution of the effect across the scalp, we also performed a quadrant analysis. After excluding midline electrodes, we grouped the remaining 44 electrodes into four quadrants of 11; left anterior, right anterior, left posterior, and right posterior (Figure 1):

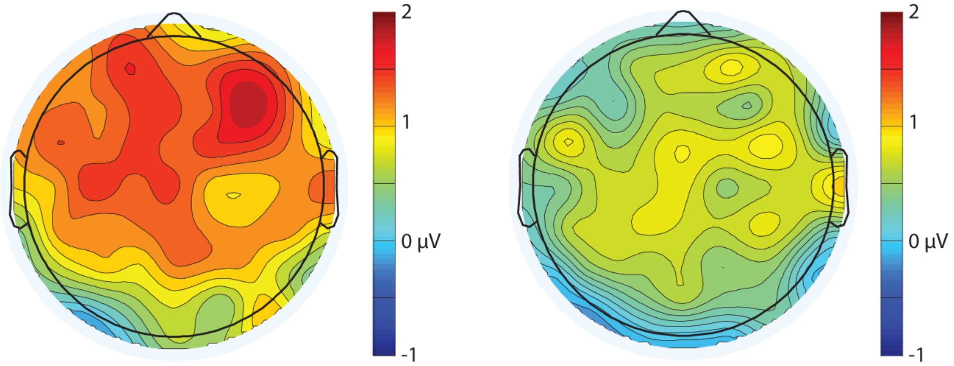


Lockwood & Tuomainen (2015) *Frontiers in Psychology* — doi.org/10.3389/fpsyg.2015.00933

Figure 1: ERPs in response to iconic and arbitrary words

A repeated measures 2x2x4 ANOVA (iconicity, sense, and quadrant) across the same 252-256ms window showed that the significant factors were iconicity (F=11.357, df=1,20, p=0.003, partial η^2=0.362) and quadrant (F=34.287, df=1,20, p<0.001, partial η^2=0.632). However, there was no significant interaction between iconicity and quadrant, which suggests that no generalizations can be made about the scalp topography of the effect of iconicity from the ANOVA. The actual difference in ERPs between conditions was approximately 1 µV at frontal sites. Waveforms and topographic plots suggested that the effect was stronger in the anterior quadrants. As predicted, there was no significant main effect of sense.

In the N400 analysis, there was a significant difference between ERP amplitudes elicited by iconic and arbitrary conditions across the time window of 350-550ms ($F=7.566$, $df=1,20$, $p=0.012$, $\eta^2=0.274$). Again, there were no other significant main effects or significant interactions. However, while this was technically significant, the waveforms in Figure 1 did not resemble a canonical N400 effect or N400 effects found in the literature. As the long-lasting increased positivity appeared to begin at around 400ms and last until the end of the trial, we also tested the time window of 400-800ms. This too was significant ($F=5.351$, $df=1,20$, $p=0.031$, $\eta^2=0.211$), and there were no significant interactions. The two main effects — the P2 at 252-256ms and the late positivity at 400-800ms — can be seen in the ERPs and topoplots in Figure 2:



252-256ms:
Iconic - arbitrary difference wave

400-800ms:
Iconic - arbitrary difference wave

Lockwood & Tuomainen (2015) *Frontiers in Psychology* — doi.org/10.3389/fpsyg.2015.00933

Figure 2: topographic plots in response to iconic and arbitrary words

Finally, if the effect was simply because of the reduplication, in that there was a visual and/or orthographic difference between the ideophones and the arbitrary words, then non-Japanese reading participants should also be sensitive to the visual reduplication. We conducted a control experiment with 34 native Dutch speakers aged 19 to 27 (24f, 10m) to establish whether the effect was due to the reduplicated nature of the ideophones. Participants saw the adverbs across the two conditions while performing an N-back task. There was no effect either at the P2 ($F=0.28$, $p=0.60$) or late positive complex ($F=0.92$, $p=0.34$) windows.

Discussion

The present study provides additional electrophysiological evidence that sound-symbolism in natural language modulates the P2 response, which in this study is the visual P2. Participants were presented with short Japanese sentences one word at a time. They were instructed to make sensibility judgments on the stimuli, unaware that the experiment was in fact designed to measure their responses to sound-symbolic ideophones and arbitrary adverbs. The ERP data showed that participants' sensitivity to the ideophone manipulation was not affected by the sensibility judgment task, and revealed different brain signatures elicited by sound-symbolic and non-sound-symbolic conditions in two windows; the P2 with a peak at 254ms, and the late effect between 400ms and 800ms. This is in line with some previous ERP studies on cross-modal congruency (Bien et al. 2012, Asano et al. 2015), but not others (Kovic et al., 2010). Since the experimental literature on sound-symbolism is still relatively sparse and reports a variety of effects, our findings are an interesting addition to the existing literature, but any theoretical claims we make based upon them can only be exploratory. We have found that ideophones elicit a greater P2 response than arbitrary adverbs, and that there is a long-lasting late effect which we take to be a late positive complex in response to the ideophones (rather than a late negative complex or elongated N400 in response to the arbitrary adverbs). We theorise that the distinctive phonology of ideophones triggers a sensory integration process of sound and the sensory information which the ideophone activates by association, which elicits the larger P2 effect, while the late positive complex may reflect the more effortful retrieval of ideophones in comparison to arbitrary words. However, word imageability and concreteness effects may also be part of this process.

The visual P2 is related to auditory input, phonological processing, and multisensory integration (Bien et al., 2012; Dien, 2009). Both of these are important factors when it comes to analysing the processing of ideophones, as ideophones have distinctive phonological structures (Bodomo, 2006; Dingemanse, 2012; Hamano, 1998) and typically depict sensory imagery in a quasi-synaesthetic way (Dingemanse, 2012; Hinton et al., 2006; Kita, 1997), in that speakers are aware that the form of the word naturally matches the word's meaning. In this experiment, it is plausible that the combination of phonological processing and a sensory integration process between sound and the triggered sensory information contributes to the heightened P2 here. The ideophones' phonological salience alone is unlikely to be the sole driver of the effect. The most obvious difference between the ideophones and the arbitrary adverbs used in this experiment was the reduplicated structure, but this also appears (albeit non-productively) in common arbitrary adverbs. It is similarly unlikely that the main factor is the individual segmental sound-symbolic properties of ideophones (Hamano, 1998), as this cannot explain the absence of the same effect for arbitrary words which use the same segments. Moreover, a control experiment with non-Japanese speaking Dutch participants ruled out any strictly visual and/or orthographic contribution to the effect. Rather, we argue that the reduplication and the specific position of certain segments, as well as the special prosodic foregrounding often assigned to ideophones (Dingemanse, 2013), results in a more holistic distinctive phonology for ideophones. This allows the speaker to recognise that this is an ideophone, a special word which depicts sensory imagery and therefore needs to be processed slightly differently. This activates a second sensory system by association, and leads to an integration process of the sound of the ideophone and the sensory information conveyed by the ideophone, in a whole process which may be considered (quasi-) synaesthetic. The ideophone-synaesthetic effect is consistent with the linguistic literature on the vivid experience of ideophones, as well as the various studies by the Osaka group, where ideophones were shown to activate the sensory areas of the brain which corresponded to the ideophones' meaning. It also echoes Brang et al's (2010) point that the P2 reflects synaesthetic experience. However, we prefer to remain agnostic over whether, and if so how far, cross-modal correspondences in language fall into the spectrum of synaesthesia.

One limitation of this study is that words were presented visually, while previous literature used auditory or concurrent audio/visual stimuli. As reading automatically activates phonological representations even in skilled readers (Savill, Lindell, Booth, West, & Thierry, 2011; Ziegler & Goswami, 2005), it is not implausible to discuss phonological processing from an experiment which presented orthographic stimuli, although it is not possible to make direct comparisons with Asano et al.'s (2015) aurally presented pseudowords. Rather than claiming the same response reflecting the same processes, we claim to identify a similar response reflecting similar processes. Likewise, directly comparing linguistic stimuli in this study with non-linguistic stimuli in the Bien et al. (2012) study would be disingenuous. However, we argue that similar underlying processes are at play. Bien et al. (2012) found that congruent cross-modal mappings elicited a larger P2 response, which was diminished after using TMS on the right intraparietal area hypothesised to be responsible for multisensory integration. Bien et al. link an increased P2 response with congruent multisensory integration between two sensory modalities; we follow that interpretation in suggesting that our results represent the sensory integration between sound and coactivated sensory information.

The present study also appears to contrast with audio-visual multisensory integration processes where hearing speech and viewing congruent speech articulation results in a decreased P2 amplitude when compared to unimodal presentation of hearing speech alone (Wassenhove, Grant, & Poeppel, 2005). However, the two modalities used in typical audio-visual integration studies involve two sensory aspects of one external audio-visual object, and the two modalities directly represent the same object. On the other hand, this study and Bien et al.'s study compare two sensory aspects from two separate origins; sound and size in Bien et al. (2012), and sound and elicited-by-association sensory information in this study. It is not implausible that multisensory integration processes are different depending on whether the senses being integrated are coming from the same source or not. A further limitation of this study was that we were not able to obtain imageability and concreteness ratings for all the stimuli used in both conditions, as adverbs in general and ideophones particularly are not used in rating studies as often as nouns and verbs. Given that descriptions of ideophones in the literature refer to their vividness and intensity, it could well be that word imageability may contribute to the effects found here. Similarly, orthographic neighbourhood was not calculated and could be a confound, but this seems unlikely as orthographic neighbourhood tends to be linked to the N400 component rather than the P2 or late positive complex (Holcomb, Grainger, & O’Rourke, 2002). Finally, another potential interpretation is that the P2 effect represents a visual process. The visual P2 has also been linked to various cognitive processes involving memory in priming tasks (Freunberger et al., 2007), as well as visual features in selective attention tasks with higher amplitudes elicited by expected visual features (Federmeier & Kutas, 2002). However, this experiment was designed with no additional working memory task or visual feature detection task, the memory requirements were limited and the visual features of the adverbs were consistent across trials, and a control experiment with non-Japanese speaking Dutch participants failed to find any visual effect. Therefore, a phonological processing and sensory integration interpretation of the P2 effect here appears most likely.

One final issue to discuss concerning the P2 in this study is its scalp topography. Given the various scalp topographies found in the general P2 literature and in the cross-modal mapping literature, we made no specific predictions. We found that the effect in our study was not restricted to any particular electrodes or quadrants, but appeared to be stronger at frontal and fronto-central areas. This is somewhat inconsistent with the P2 findings of Bien et al. (2012) and the 1-300ms gamma band increase in Asano et al. (2015), who localised their effects to centro-parietal regions. We argue that this is related to how we presented the stimuli; as the visual P2 tends to be maximally frontal while the auditory P2 tends to be maximally central (Key et al., 2005), it is perhaps not surprising that this study, which presented visual words only, found an apparently more frontal effect compared to studies which presented auditory and visual stimuli. None of these three studies is consistent with Kovic et al. (2010), who found an occipital effect of sound-symbolism. Moreover, their effect was approximately 100ms earlier and was a negative-going wave. We can only speculate as to the reasons for the three-fold difference between Kovic et al. (2010) and the present study, Bien et al. (2012) and Asano et al. (2015), but this difference may arise from the fact that participants in Kovic et al. (2010) had been trained to memorise congruent and incongruent mappings, whereas the other studies relied on spontaneous judgments based on intuition or native language.

The N400 test in the 350-550ms timeframe was significant, which would at first glance echo Asano et al.'s (2015) finding that sound-symbolism facilitates semantic integration. However, with the possible exception of the waveforms in the left posterior quadrant, the response in this study does not resemble an N400 response, and so to discuss our significant findings as if they were typical N400s would be disingenuous. Rather, our results more closely resemble a late sustained positive complex, and in this section we speculate as to what this may mean. The late sustained positive complex in response to ideophones may reflect the necessity for more post-lexical processing (Lau, Phillips, & Poeppel, 2008), due to an increased processing cost of sound-symbolism in language. Ideophones in Japanese have high frame specificity and can have non-standard syntactic requirements (Akita, 2012; Dingemanse, 2012), meaning that the syntactic integration of ideophones into a sentence is perhaps harder than it is for arbitrary words. This late positive complex may be a neural representation of the trade-off between the expressivity of sound-symbolism and the efficiency of arbitrariness.

One further issue which future studies should address is the precise mechanism of the putative synaesthetic effect; why should phonologically distinct ideophones (and not, say, the phonological distinctiveness of loan words which use foreign phonology) trigger a sensory integration effect? It appears that various factors contribute to the sound-symbolism of ideophones; consonants, vowels, syllabic structure, and prosody, as well as the semantic domain of sensory imagery that they express. The combination of some or all of these factors is probably what makes ideophones special, while having just one factor is not enough to make other words sufficiently sound-symbolic. It remains to be seen whether Japanese speakers' sensitivity to these special factors is simply statistically learned (and therefore making Japanese ideophones examples of conventional sound-symbolism) or whether there is some degree of inherent sound-symbolism to these factors; the behavioural experiments on English speakers with no knowledge of Japanese statistical patterns would suggest that it is the latter (Iwasaki, Vinson, & Vigliocco, 2007a; Iwasaki et al., 2007b; Kantartzis et al., 2011; Oda, 2000; Yoshida, 2012), although there is always some inevitable conventionalism of ideophones' form within any given language. The replication of this study using ideophones in a different language would go some way towards answering this.

Conclusions

This study has shown that sound-symbolic adverbs in Japanese have a significant effect on the P2 response and in the 400-800ms timeframe when compared with arbitrary adverbs, and we speculate that this effect may be due to the distinctive phonological properties of ideophones precipitating a sensory integration process between sound of the ideophones and the sensory representations of the triggered sensory domains. Further research is needed to clarify whether this effect is generalizable to all ideophones, and whether this effect can be replicated with speakers of other languages. However, this study provides exciting evidence that sound-symbolism is not just psycholinguistically detectable in deliberately constructed pseudowords, but also in real sound-symbolism in natural language.

References

Abel, G. A., & Glinert, L. H. (2008). Chemotherapy as language: Sound symbolism in cancer medication names. *Social Science & Medicine*, *66*(8), 1863–1869.

Akita, K. (2009). A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics. Kobe University.

Akita, K. (2012). Toward a frame-semantic definition of sound-symbolic words: A collocational analysis of Japanese mimetics. *Cognitive Linguistics*, *23*(1).

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex*, *63*, 196–205.

Barber, H. A., Ben-Zvi, S., Bentin, S., & Kutas, M. (2011). Parafoveal perception during sentence reading? An ERP paradigm using rapid serial visual presentation (RSVP) with flankers. *Psychophysiology*, 48(4), 523–531.

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

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663–672.

Bodomo, A. (2006). The Structure of Ideophones in African and Asian Languages: The Case of Dagaare and Cantonese. In *Selected Proceedings of the 35th Annual Conference on African Linguistics: African Languages and Linguistics in Broad Perspectives*.

Brang, D., Kanai, S., Ramachandran, V. S., & Coulson, S. (2010). Contextual Priming in Grapheme–Color Synesthetes and Yoked Controls: 400 msec in the Life of a Synesthete. Journal of Cognitive Neuroscience, 23(7), 1681–1696.

Carpenter, A. L., & Shahin, A. J. (2013). Development of the N1-P2 auditory evoked response to amplitude rise time and rate of formant transition of speech sounds. *Neuroscience Letters*, *544*, 56–61.

Childs, G. T. (1994). African ideophones. In J. J. Ohala, L. Hinton, & J. Nichols (Eds.), *Sound Symbolism*.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *115*(4), 732–744.

Davis, R. (1961). The Fitness of Names to Drawings. a Cross-Cultural Study in Tanganyika. *British Journal of Psychology*, *52*(3), 259–268.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. Journal of Neuroscience Methods 134:9-21

Dien, J. (2009). The neurocognitive basis of reading single words as seen through early latency ERPs: A model of converging pathways. *Biological Psychology*, *80*(1), 10–22.

Diffloth, G. (1972). Notes on expressive meaning. *Chicago Linguistic Society*, *8*, 440–447.

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, *6*(10), 654–672.

Dingemanse, M. (2013). Ideophones and gesture in everyday speech. *Gesture*, *13*(2), 143–165.

Doke C. M. (1935). Bantu Linguistic Terminology. London: Longmans, Green, and Co.

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730–747.

Firth, J. R. (1964). *The tongues of men, and Speech*. Oxford University Press.

Freunberger, R., Klimesch, W., Doppelmayr, M., & Höller, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters*, *426*(3), 181–186.

Gasser, M., Sethuraman, N., & Hockema, S. (2010). Iconicity in expressives: an empirical investigation. In S. Rice & J. Newman (Eds.), *Experimental and Empirical Methods in the Study of Conceptual Structure, Discourse, and Language* (pp. 163–180). Stanford, CA: CSLI Publications.

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, *11*(5), 473–490.

Hamano, S. (1998). *The Sound-Symbolic System of Japanese*. Center for the Study of Language and Information ; Tokyo.

Hinton, L., Nichols, J., & Ohala, J. J. (2006). *Sound Symbolism*. Cambridge University Press.

Holcomb, P., Grainger, J., & O'Rourke, T. (2002). An Electrophysiological Study of the Effects of Orthographic Neighborhood Size on Printed Word Perception. Journal of Cognitive Neuroscience, 14(6), 938–950.

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

Ishiguro, H. (1993) Onomatope no hassei. *Gengo*, 22(6), 26-33.

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007a). How does it hurt, kiri-kiri or siku-siku? Japanese mimetic words of pain perceived by Japanese speakers and English speakers. In M. Minami (Ed.), *Applying theory and research to learning Japanese as a foreign language* (pp. 2–19). Newcastle: Cambridge Scholars Publishing.

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007b). What do English speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education Around the Globe*, *17*, 53–78.

Johnson, R. C. (1967). Magnitude symbolism of English words. *Journal of Verbal Learning and Verbal Behavior*, *6*(4), 508–511.

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How Sound Symbolism Is Processed in the Brain: A Study on Japanese Mimetic Words. *PLoS ONE*, *9*(5), e97905.

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

Key, A. P. F., Dove, G. O., & Maguire, M. J. (2005). Linking Brainwaves to the Brain: An ERP Primer. *Developmental Neuropsychology*, *27*(2), 183–215.

Kilgarriff, A., Rychly, P., Smrz, P., Tugwell, D. The Sketch Engine. Proc EURALEX 2004, Lorient, France; Pp 105-116. http://www.sketchengine.co.uk (last accessed: 06/09/13)

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics*, *35*(2).

Köhler, W. (1947). *Gestalt psychology; an introduction to new concepts in modern psychology (Rev. ed.)*. Oxford, England: Liveright.

Kong, L., Zhang, J. X., Kang, C., Du, Y., Zhang, B., & Wang, S. (2010). P200 and phonological processing in Chinese word recognition. *Neuroscience Letters*, *473*(1), 37–41.

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

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

Kutas, M., DeLong, K., & Kiang, M. (2010). From phonemes to discourse: Event-related brain componentry and paradigms for investigating normal and abnormal language processing. In J. Guendozi, F. Loncke, & M. J. Williams (Eds.), *The Handbook of*

Psycholinguistic and Cognitive Processes: Perspectives in Communication Disorders (pp. 119–144).

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, *102*(1), 30–45.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT.

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and Weak. *Current Directions in Psychological Science*, *10*(2), 61–65.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound–shape correspondences in toddlers and adults. *Developmental Science*, *9*(3), 316–322.

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory Visual–Auditory Object Recognition in Humans: a High-density Electrical Mapping Study. *Cerebral Cortex*, *14*(4), 452–465.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128.

Oda, H. (2000). *An embodied semantic mechanism for mimetic words in Japanese*. Indiana University, Indiana.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice. In L. Hinton, J. J. Ohala, & J. Nichols (Eds.), *Sound symbolism* (pp. 325–347). New York: Cambridge University Press.

Osaka, N. (2009). Walk-related mimic word activates the extrastriate visual cortex in the human brain: an fMRI study. *Behavioural Brain Research*, *198*(1), 186–189.

Osaka, N. (2011). Ideomotor response and the neural representation of implied crying in the human brain: An fMRI study using onomatopoeia1. *Japanese Psychological Research*, *53*(4), 372–378.

Osaka, N., & Osaka, M. (2005). Striatal reward areas activated by implicit laughter induced by mimic words in humans: a functional magnetic resonance imaging study. *Neuroreport*, *16*(15), 1621–1624.

Osaka, N., & Osaka, M. (2009). Gaze-related mimic word activates the frontal eye field and related network in the human brain: an fMRI study. *Neuroscience Letters*, *461*(2), 65–68.

Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., & Shibasaki, H. (2003). An emotion-based facial expression word activates laughter module in the human brain: a functional magnetic resonance imaging study. *Neuroscience Letters*, *340*(2), 127–130.

Osaka, N., Osaka, M., Morishita, M., Kondo, H., & Fukuyama, H. (2004). A word expressing affective pain activates the anterior cingulate cortex in the human brain: an fMRI study. *Behavioural Brain Research*, *153*(1), 123–127.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, *1*(E227).

Ramachandran V.S., & Hubbard E.M. (2001). Synaesthesia -- A window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. *Brain and Language*, *128*(1), 18–24.

Saussure, F. de. (1959). *Course in general linguistics*. New York: Philosophical Library.

Savill, N., Lindell, A., Booth, A., West, G., & Thierry, G. (2011). Literate humans sound out words during silent reading: *NeuroReport*, 22(3), 116–120. doi:10.1097/WNR.0b013e328342ed27

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, *16*(16), 1781–1785.

Shogakukan. (2001). *Shogakukan Progressive English-Japanese Dictionary* (3rd edition.). Shogakukan.

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, *39*(4), 553–569.

Wassenhove, V. van, Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(4), 1181–1186.

Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., & Schröger, E. (2007). Binding Symbols and Sounds: Evidence from Event-Related Oscillatory Gamma-Band Activity. *Cerebral Cortex*, *17*(11), 2696–2702.

Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schröger, E. (2004). From symbols to sounds: Visual symbolic information activates sound representations. *Psychophysiology*, *41*(5), 709–715.

Yoshida, H. (2012). A Cross-Linguistic Study of Sound Symbolism in Children's Verb Learning. *Journal of Cognition and Development*, *13*(2), 232–265.

Yuval-Greenberg, S., & Deouell, L. Y. (2007). What You See Is Not (Always) What You Hear: Induced Gamma Band Responses Reflect Cross-Modal Interactions in Familiar Object Recognition. *The Journal of Neuroscience*, *27*(5), 1090–1096.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3–29. doi:10.1037/0033-2909.131.1.3

http://www.biosemi.com/headcap.htm (last accessed: 21/07/14)

Chapter 4

Sound symbolism boosts novel word learning.

Published as:

Lockwood, Dingemanse, & Hagoort (2016). Sound-Symbolism Boosts Novel Word Learning. Journal of Experimental Psychology. Learning, Memory, and Cognition.

Abstract

The existence of sound symbolism (or a non-arbitrary link between form and meaning) is well-attested. However, sound symbolism has mostly been investigated with pseudowords in forced choice tasks, neither of which are representative of natural language. This study uses ideophones, which are naturally occurring sound-symbolic words that depict sensory information, to investigate how sensitive Dutch speakers are to sound symbolism in Japanese in a learning task. Participants were taught two sets of Japanese ideophones; one set with the ideophones' real meanings in Dutch, the other set with their opposite meanings. In Experiment 1, participants learned the ideophones and their real meanings much better than the ideophones with their opposite meanings. Moreover, despite the learning rounds, participants were still able to guess the real meanings of the ideophones in a two-alternative forced choice (2AFC) test after they were informed of the manipulation. This shows that natural language sound symbolism is robust beyond 2AFC paradigms and affects broader language processes such as word learning. In Experiment 2, participants learned regular Japanese adjectives with the same manipulation, and there was no difference between real and opposite conditions. This shows that natural language sound symbolism is especially strong in ideophones, and that people learn words better when form and meaning match.

Introduction

The classical view that the relation between the sound and meaning of lexical items is arbitrary is at odds with growing evidence from sound symbolism research (Lockwood & Dingemanse, 2015; Perniss, Thompson, & Vigliocco, 2010). Sound symbolism has mostly been investigated with variations on the *bouba/kiki* paradigm, where participants associate spiky shapes and round shapes with the pseudowords *kiki* and *bouba* respectively in a forced choice task. Original forced choice experiments established that regardless of language background, people associate specific sounds with specific sensory properties, such as object roundness with vowel roundness and object size with vowel height and backness (Davis, 1961; Köhler, 1947; Newman, 1933; Ramachandran & Hubbard, 2001; Sapir, 1929). More recent behavioural experiments have probed how these effects vary when accounting for individual vowels and consonants (Nielsen & Rendall, 2013), increasing the number of choices available (Aveyard, 2012), setting the experiment up as a gradient of choice rather than alternative choices (Thompson & Estes, 2011), and replicating the paradigm with non-WEIRD (Western, educated, industrialised, rich and democratic) (Henrich, Heine, & Norenzayan, 2010) participant groups to examine the effects of culture, orthography, and specific neurological condition (Bremner et al., 2013; Drijvers, Zaadnoordijk, & Dingemanse, 2015; Occelli, Esposito, Venuti, Arduino, & Zampini, 2013).

These studies have been instrumental in establishing that people can reliably make certain sound-meaning associations. However, the stimuli used in these experiments are pseudowords which, in most cases, are deliberately constructed to maximise contrasts. This does not guarantee that sound symbolism research based on these experiments is representative of sound symbolism in natural language, and therefore may not directly address the processes at play in natural language learning and use.

Sound symbolism in natural languages is more common than often assumed (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Hinton, Nichols, & Ohala, 1994), and offers many possibilities for experimental investigation. Most strikingly, many of the world's languages—though not the small set of Western languages commonly studied in psycholinguistics (Majid & Levinson, 2010)—feature sizable classes of ideophones, sound-symbolic words with vivid sensory meanings (Voeltz & Kilian-Hatz, 2001; Dingemanse, 2012). Japanese, for instance, has thousands of these words (also known as 'mimetics'), and their phonological, morphological, syntactic, and semantic properties are well documented (Akita, 2009; Hamano, 1998; Shogakukan, 2001). Sound symbolism research which uses ideophones has broadly confirmed the effects from pseudoword experiments, showing that participants are able to guess the meaning of sound-symbolic real words at above-chance accuracy (Iwasaki, Vinson, & Vigliocco, 2007a, 2007b; Nygaard, Herold, & Namy, 2009; Oda, 2000; Revill, Namy, DeFife, & Nygaard, 2014). While two-alternative forced choice (2AFC) experiments with ideophones have replicated findings from the pseudoword literature using natural language stimuli, the restricted nature of such tasks does not allow strong inferences about the strength of sound symbolism in ideophones beyond the 2AFC context, or about the functions of sound symbolism in language learning.

Developmental studies have shown that natural sound symbolism scaffolds word learning in infants (Imai, Kita, Nagumo, & Okada, 2008[1]; Laing, 2014; Yoshida, 2012), perhaps by facilitating multisensory integration of linguistic form and the sensory properties of the referent (Asano et al., 2015). Previous experiments on word learning in adults have shown that adults also look for sound-symbolic cues (Nygaard, Cook, & Namy, 2009), and this suggests that natural sound symbolism can be identified outside a 2AFC paradigm and exploited during word learning.

Nygaard et al. used a set of real words in a sound symbolism experiment, which is a welcome departure from the typical pseudoword 2AFC experiments which tend to dominate the sound symbolism literature (Lockwood & Dingemanse, 2015). The words they used are mostly adjectives, some verbs, and even a couple of nouns—all categories considered to be arbitrary in the Japanese linguistic literature. They found judgement accuracy and reaction time differences between words learned with their real translations and with random translations, with participants responding to words learned with their real translations more accurately and more quickly. They use this to support their argument that sound-symbolically congruent mappings between sound and word meaning facilitate word learning. However, they found no difference in either judgement accuracy or reaction time between words learned with their real and opposite translations.

While there may well be sound-symbolic traces or informative prosodic contours in the words (Kunihira, 1971), the nature of the sound-symbolic links is uncertain, and may be best described as *covert* sound symbolism. Ideophones, on the other hand, are an example of *overt* sound symbolism. Ideophones stand out from other words due to their morphophonological patterns, and work in linguistics and psychology has found consistent links between the sounds of the ideophones and their cross-modally congruent meanings (Vigliocco & Kita, 2006; Dingemanse et al., 2015). A learning study based on overtly sound-symbolic ideophones rather than covertly sound-symbolic regular words can clarify the role of sound symbolism in language learning.

This study was designed to investigate whether adult participants learn words better when form and meaning match. We designed a learning and recognition experiment where Dutch participants learned Japanese ideophones with either their real translation (i.e. where the Japanese ideophones are sound-symbolically congruent with the Dutch translations) or their opposite translation (i.e. where the Japanese ideophones are sound-symbolically incongruent with the Dutch translations). We hypothesised that participants would learn the real translations better than the

---

[1] While Imai et al. (2008) used non-words for this experiment, they did so based on standard Japanese ideophonic templates, which were then approved as acceptable, naturalistic words by adults. The use of non-words in Imai et al. (2008) is far more like the use of real ideophones than the use of *bouba/kiki*-esque materials.

opposite translations because for real translations, sound-symbolic cues in ideophones would highlight perceptual analogies between form and meaning and thereby facilitate learning. We also tested the participants on a 2AFC task afterwards because we hypothesised that participants would still be sensitive to the sound-symbolic cues in ideophones despite the learning task. If participants are unable to guess the meanings of the ideophones in a 2AFC task afterwards (or if participants just selected the options which they learned earlier), then learning a word can suppress participants' sensitivity to sound-symbolic cues; if participants were able to disregard their earlier learned associations and guess the real meanings of the words at above chance level, then sound-symbolic cues are still available to participants despite learning a specific word-to-word mapping. Finally, we hypothesised that in a control experiment with regular adjectives, there would either be a much smaller learning effect or no learning effect at all when learning the real translations compared to the opposite translations, because overt sound-symbolic cues are not available.

Material and methods

Stimuli selection

First pre-test

We made a list of 376 reduplicated CVCV-CVCV Japanese ideophones, and translated a systematic selection of them in Dutch. Translations were agreed upon by GL, MD (a native Dutch speaker and ideophone expert), and a native Dutch speaker who is fluent in Japanese. We filtered out ideophones which had strongly similar forms and meanings (e.g. *bakibaki* and *bokiboki*, both of which mean a cracking sound like of tree branches or knuckles), and kept the most frequent or canonical ideophone only. We also filtered out ideophones where a simple Dutch translation couldn't fully distinguish between different concepts (e.g. *hatahata*, *batabata*, and *patapata*, all of which mean "flapping" but to a greater or lesser degree). Finally, we aimed for translations that were as short and as uniform across opposites as possible, so we filtered out ideophones where it was not possible to get a good translation of its opposite meaning (e.g., we could not find an opposite to *muzumuzu*, meaning "itchy", other than "not itchy").

This left us with 95 ideophones which had good Dutch translations for both their real and opposite meanings. We used the CELEX database to ensure that there was no difference in word frequency between the real and opposite translations. There were also no differences between conditions in terms of word length and the number of letters in common between the translation and the ideophone. We recorded a female native Japanese speaker, unaware of the experimental manipulation, reading aloud each ideophone in a soundproof booth. Recordings were then checked with another native Japanese speaker to ensure that intonation and pitch accent were natural.

We conducted a stimuli selection pre-test with 26 native Dutch speakers (9m, 17f, 22-35 years old) in order to see whether participants could guess the meaning of the ideophone at above chance levels. This was a 2AFC task, where participants saw and heard the ideophone, then saw two possible Dutch translations; the real translation and the opposite translation (although 2AFC tasks have their limitations—as we point out above and elsewhere (Lockwood & Dingemanse, 2015)—we find they can be useful when supplemented with other methods, as here). Participants were instructed to pick the translation which best matched the ideophone by pressing the left CTRL key to select the word on the left, and the right CTRL key to select the word on the right. This is illustrated in Figure 1.

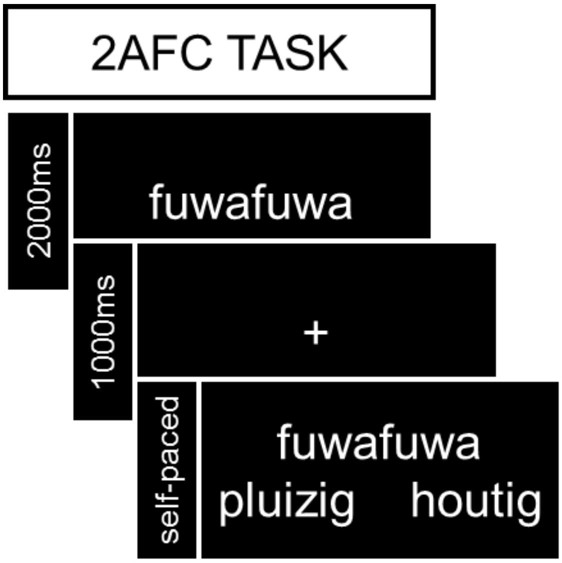


Figure 1: stimuli selection pre-test procedure

We used Presentation software to present stimuli and record responses. Participants guessed the ideophones correctly 63.1% of the time, which was above chance (95% CIs: 60.6% - 65.7%, $\mu = 0.5$, $p < 0.001$). Even though this was a stimuli selection pre-test, it is of interest to show that people with no knowledge of Japanese can guess the meanings of a large selection of ideophones at above chance accuracy. This is generally taken for granted, but our study is, to our knowledge, the most extensive demonstration of this beyond Iwasaki et al. (2007b), who only tested two semantic domains by using 24 mimetic words for laughing and 28 mimetic words for walking.

Afterwards, we asked participants whether there were any ideophones which resembled related Dutch words. This filtered out confound words like *wakuwaku* (meaning "excited", but also a Dutch children's TV show), *iraira* (meaning "angry", but too close to English words such as *irate* and *irritated*), and *pikapika* (meaning "bright, flashing", but also the battle cry of Pikachu, a character who attacks using flashes of electricity (Oak, 1996), from the Pokémon video game and TV show which was popular with participants of this generation). We selected the 50 ideophones which were guessed most accurately in the pretest (which was over 63% of the time). The entire 2AFC test showed that ideophones are, on the whole, sound-symbolically informative to Dutch speakers, but to home in on potential learning effects, we used the individual ideophones which were most obviously sound-symbolic.

We removed 12 ideophones from these 50; *pikapika*, *iraira*, and *wakuwaku* due to world knowledge confounds, four others due to the fact that one of the translations shared the same first letter as the ideophone, *suyasuya*, meaning "sleeping peacefully", which we could not find a one word translation for, and four more ideophones which were guessed at under 50% accuracy in the second pre-test which was used to pilot the learning task.

Second pre-test

This second pre-test was actually intended as the main experiment, and involved participants learning the ideophones by making 2AFC decisions and then receiving feedback about whether they were correct. Participants saw and heard an ideophone, and then saw two Dutch translations; the real translation and the opposite translation. When they selected one, they were informed whether they were "correct" (i.e. if they had chosen the real translation in the real condition, or the opposite translation in the opposite condition) or "incorrect". This continued for three rounds or until participants could choose the correct word over 80% of the time (which occasionally took four or five rounds). They then performed one final 2AFC test. We hypothesised that participants would find it harder to remember the Dutch translations for ideophones in the opposite condition, and this is indeed what we found; participants made significantly more mistakes in the final 2AFC test when choosing the translations for the ideophones in the opposite condition than ideophones in the real condition. However, there was a major confound: we used the same real and opposite translations in each learning round, which meant that about a third of the participants realised that they could ignore the ideophone and just remember which of two Dutch words to choose each time. This is the point at which we decided that moving beyond 2AFC experiments was essential. Despite this, though, the participants still made more mistakes in the opposite condition. The participants in this second pre-test were divided into two groups where the ideophones in real and opposite conditions were counterbalanced. Participants in Group 1 made an average of 9.33 mistakes in the real condition and 14.07 mistakes in the opposite condition; participants in Group 2 made an average of 9.4 mistakes in the real condition and 15.33 mistakes in the incorrect condition. As both groups recalled the real translations better than the opposite translations, we did not counterbalance the ideophones across conditions in the full experiment; each participant learned half the ideophones in the real condition and half the ideophones in the opposite condition, and these were the same across participants.

Experiment procedure: Experiment 1

We used the 38 ideophones from the pre-test for Experiment 1, where we tested 32 participants (10m, 22f). As in the pre-test, there were no differences in the number of letters in common between the ideophones and between the Dutch words across conditions. We used the CELEX database to additionally ensure that there was no difference in word frequency between the Dutch words which the participants learned in the real and opposite conditions. Two participants were discarded; one for pressing the wrong response buttons throughout the experiment, the other for taking an abnormally long time during the self-paced learning sessions (reaction times were not recorded for this part, but lab notes taken at the time noted that the participant took a lot longer to complete the task). This resulted in 30 participants whose data we analysed.

Participants learned the real translations to 19 ideophones and the opposite translations to the other 19 ideophones. In one learning round, participants saw each ideophone and translation once, and then saw the ideophone and translation together. There were two learning rounds in total. The order of Dutch words and ideophones was randomised for each round and for each participant, but the items and conditions were fixed across participants.

REAL condition		OPPOSITE condition	
ideophone	translation	ideophone	translation
fuwafuwa ("fluffy")	pluizig ("fluffy")	kibikibi ("energetic")	futloos ("tame, tired")
boroboro ("worn out")	versleten ("worn out")	ukiuki ("happy")	verdrietig ("sad")

Table 1: two example stimuli for each condition

In the learning round, the initial Dutch word was presented for 1000ms with 100ms of jitter, followed by a fixation cross for 1000ms with 100ms of jitter. As the ideophone was played over the speakers, the ideophone was presented visually for 2000ms with 200ms of jitter. This was again followed by a fixation cross. The final screen with the ideophone and its Dutch meaning was presented until participants were happy to move onto the next item. Between trials, a blank screen was presented, followed by a fixation cross to announce the beginning of the next trial.

In the test round, participants were presented with either the word pairs that they had learned, or a pseudo-randomised pairing of ideophones and translations which they had not seen paired together before. These pairings were pseudo-randomised to ensure that the meanings were semantically unrelated (for example, the Japanese *fuwafuwa*, learned as "fluffy", and the Dutch *kortaf*, meaning "curt"). Participants were instructed to answer Yes (indicating that this was a word pair that they had learned) or No (indicating that this was not a word pair that they had learned) using the left CTRL key for Yes and the right CTRL key for No. Pairs requiring a Yes response made up 50% of the trials. As in the learning round, participants saw the Dutch word first, then saw and heard the Japanese ideophone, but this time they were asked to respond as soon as possible after seeing and hearing the Japanese ideophone rather than waiting for a screen where both words were presented at the same time. Timings in the test stage were identical to the learning stage. The fixation cross was displayed until participants responded, at which point a blank screen was presented, followed by a fixation cross to announce the beginning of the next trial.

After the test round, we told the participants that half the words they had learned had the real meanings, but half actually had the opposite meaning. We asked them to forget everything they had learned, and instead to choose which translation they felt was more natural for each ideophone. Participants saw and heard the ideophone, and then saw two possible Dutch translations; the real one and the opposite one (i.e., they saw the translation they had learned, and the opposite of that translation). They selected what they felt was the most natural translation by pressing the left CTRL key for the translation on the left and the right CTRL key for the translation on the right. As in the pre-test, there were no differences between the frequencies of the real and opposite Dutch words. Timings were identical to the earlier stages. The final screen was displayed until participants responded. This was identical to the stimuli selection pre-test, and the full procedure is illustrated in Figure 2.

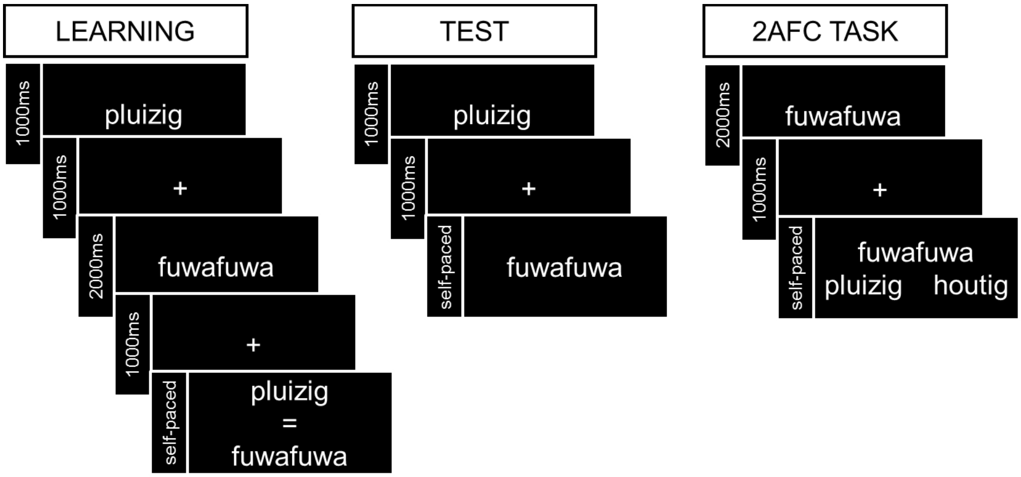


Figure 2: post-test sound-symbolic sensitivity check procedure

Experiment 2: stimuli selection and experiment procedure

We ran a second experiment with regular adjectives —i.e., presumably non-sound-symbolic words— to investigate to what extent behavioural effects found were due to the sound-symbolic nature of ideophones. This experiment was done with a separate group of 30 participants.

Stimuli were selected in a similar way to Experiment 1. We created a list of 87 Japanese adjectives and translated them into Dutch for both their real and opposite meanings. We used the CELEX database to ensure that there was no difference in word frequency between the real and opposite translations, nor was there a difference in word frequency between the words which the participants learned in real and opposite conditions. There were also no differences between conditions in terms of word length and the number of letters in common between the translation and the regular adjective. The same female native Japanese speaker provided the recordings.

We conducted a stimuli selection pre-test with 28 native Dutch speakers (9m, 17f, 20-40 years old) in order to see whether participants could guess the meaning of the words at above chance levels. The procedure was identical to the first pre-test for Experiment 1. We used Presentation software to present stimuli and record responses. Participants guessed the words correctly 55.3% of the time, which was above chance (95% CIs: 53.4% — 57.2%, $\mu = 0.5$, $p < 0.001$). We asked participants afterwards whether there were any ideophones which resembled related Dutch words. The word *kawaii* (meaning "cute") was filtered out, because this is a well-known word in popular culture. We also excluded words which were shorter than three syllables to keep the Japanese word length consistent across the two experiments. We selected the 38 most correctly guessed regular adjectives in order to remain consistent with Experiment 1. All were guessed above 53.6% (average 55.3%).

In Experiment 2, we tested 30 participants (8m, 22f), and the procedure was exactly the same as in Experiment 1.

Results

Experiment 1: Ideophone learning

Participants made more recognition mistakes in the opposite condition than in the correct condition; participants correctly remembered the real word pairing 86.1% of the time, but correctly remembered the opposite word pairing only 71.1% of the time (Figure 3).

As the dependent variable was binary —correct or incorrect— we analysed the responses using a mixed-effects logit model with the glmer function of the lme4 (versions 1.1-8) package in R. The data was modelled by including a per-participant and per-item random adjustment to the fixed intercept with a condition random slope for the fixed effect by participant. The condition was sum contrast coded.

Model comparison showed that a random effect by ideophone did explain some variance in the data (log likelihood difference = 7.12, $\chi^2 = 14.24$, df = 1, $p < 0.001$). That means that some ideophones were answered correctly more often than others. However, even when controlling for this random effect by ideophone, model comparison still showed a significant fixed effect of condition ($\beta = -0.5978$, log likelihood difference = 8.85, $\chi^2 = 17.695$, df = 1, $p < 0.001$). The model estimated that ideophones learned in the real condition were answered 9.53 percentage points more accurately than ideophones learned in the opposite condition.

There were also significant differences in reaction times between conditions, with participants responding faster to ideophones in the real condition (mean RT = 1794ms) than the opposite condition (mean RT=2280ms). The data was modelled by including a per-participant and per-item random adjustment to the fixed intercept with a condition random slope for the fixed effect by participant. The condition was sum contrast coded. The model showed a significant fixed effect of condition ($\chi^2 = 13.92$, $p < 0.001$). This difference existed even when only analysing correctly answered trials ($\chi^2 = 8.10$, $p = 0.0044$), and so is not just a speed/accuracy trade off. There was also a strong correlation between the number of correct responses per ideophone and the speed of the reaction to that ideophone; the better an ideophone was remembered, the faster it was responded to ($r = -0.67$, $p < 0.001$). However, there was no correlation between the number of correct responses per participant and reaction times, meaning that more accurate participants were not necessarily faster at responding.

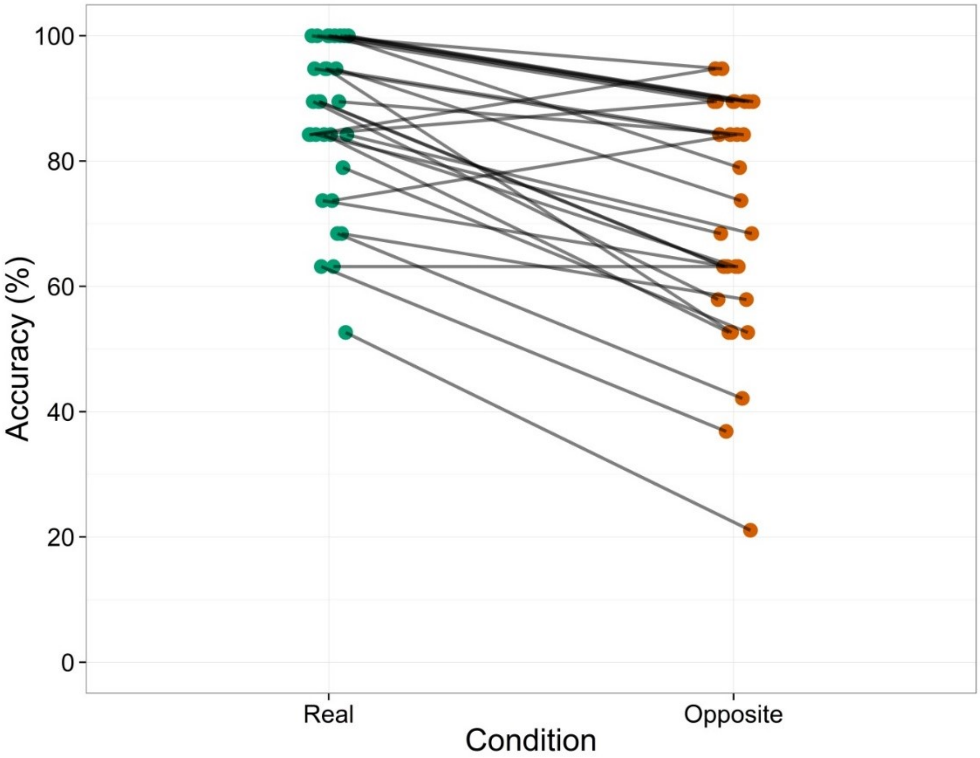


Figure 3: ideophone recognition accuracy per condition. Dots represent a participant's performance in each condition, lines connect the same participant across conditions.

In the sound symbolism sensitivity check after the experiment, participants guessed the real meanings of the Japanese words with 72.3% accuracy, which was comfortably above chance ($\mu = 0.5$, $t = 10.51$, $df = 29$, $p < 0.001$). Only one participant guessed the meanings at below 50% accuracy, one ideophone (*tsuyatsuya*) was guessed at 50% accuracy, and only one ideophone (*gowagowa*) was guessed at below 50% accuracy (although it was guessed at 69% accuracy during the pre-test). We checked if participants who guessed more accurately also guessed faster, but there was no link between reaction times and accuracy ($r = 0.13$, $n = 30$, $p = 0.49$). We also checked if ideophones which were guessed more accurately were also guessed more quickly. There was a correlation between reaction time and the mean accuracy at which the ideophone was guessed ($r = -0.39$, $n = 38$, $p = 0.015$), and this was not a speed/accuracy trade off.

Experiment 2: Regular adjective learning

In contrast to Experiment 1, there was no learning effect present in Experiment 2, which used regular adjectives instead of ideophones: participants correctly remembered the real word pairing 79.1% of the time, and the opposite word pairing 77% of the time (Figure 4).

We analysed the responses using the same mixed-effects logit model and modelling procedure as in Experiment 1. Model comparison showed a random effect by regular adjective did explain some variance in the data (log likelihood difference = 12.86, $\chi^2 = 25.718$, $df = 1$, $p < 0.001$). That means that some regular adjectives were answered correctly more often than others. However, when controlling for this random effect by regular adjective, model comparison showed no fixed effect of condition ($\beta = -0.1256$, log likelihood difference = 0.38, $\chi^2 = 0.7739$, $df = 1$, $p = 0.379$). The model estimated that regular adjectives learned in the real condition were answered 1.81 percentage points more accurately than regular adjectives learned in the opposite condition.

Similarly, there were no statistical differences in reaction times between the two conditions, either for all trials ($\chi^2 = 0.14$, $p = 0.70$) or correctly answered trials only ($\chi^2 = 0.51$, $p = 0.48$). Both conditions in both experiments are compared in Figure 5. In the sound symbolism check after the experiment, participants guessed the real meanings of the regular adjectives with 63% accuracy, which was again above chance ($\mu = 0.5$, $t = 7.21$, $df = 29$, $p < 0.001$). This is far lower than the 72.3% accuracy in the ideophone condition, but these figures cannot be compared directly as it involves two different groups of participants guessing the meanings of two different sets of words. Three participants guessed the meanings at below 50% accuracy, one word was guessed at exactly 50% accuracy, and four words were guessed at below 50% accuracy. For measures involving reaction times, the correlations were reversed relative to Experiment 1. There was no correlation between the number of correct responses per ideophone and the speed with which participants responded to them ($r = -0.1$, $p = 0.54$), but there was a correlation for participants' accuracy and reaction times, in that the more accurate participants took longer to guess the words (r = 0.46, p = 0.011).

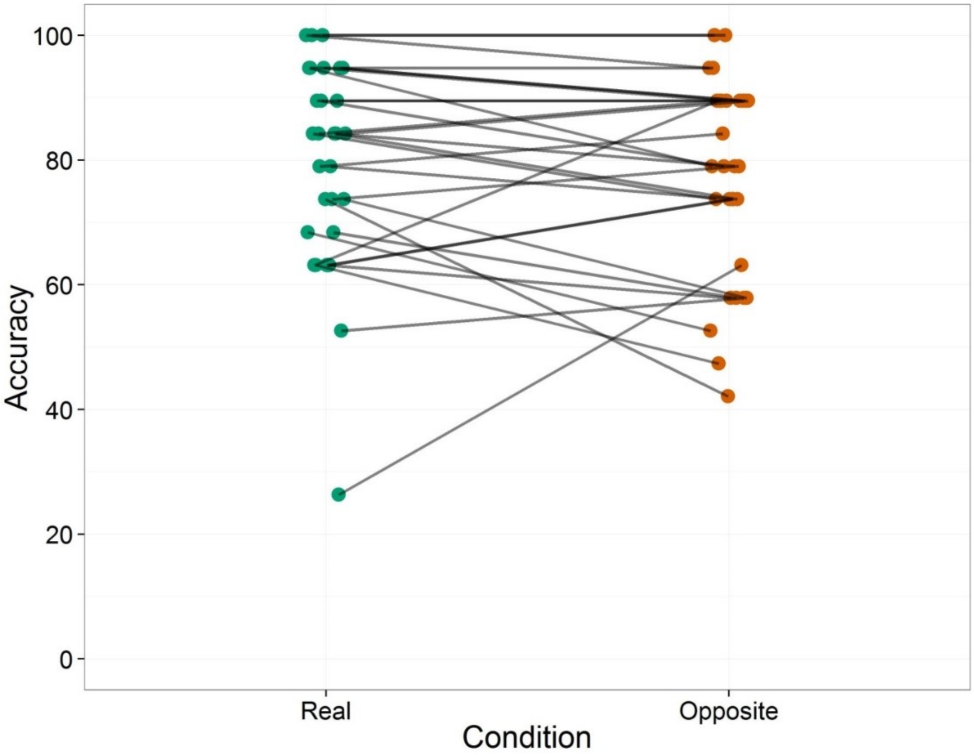


Figure 4: regular adjective recognition accuracy per condition. Dots represent a participant's performance in each condition, lines connect the same participant across conditions.

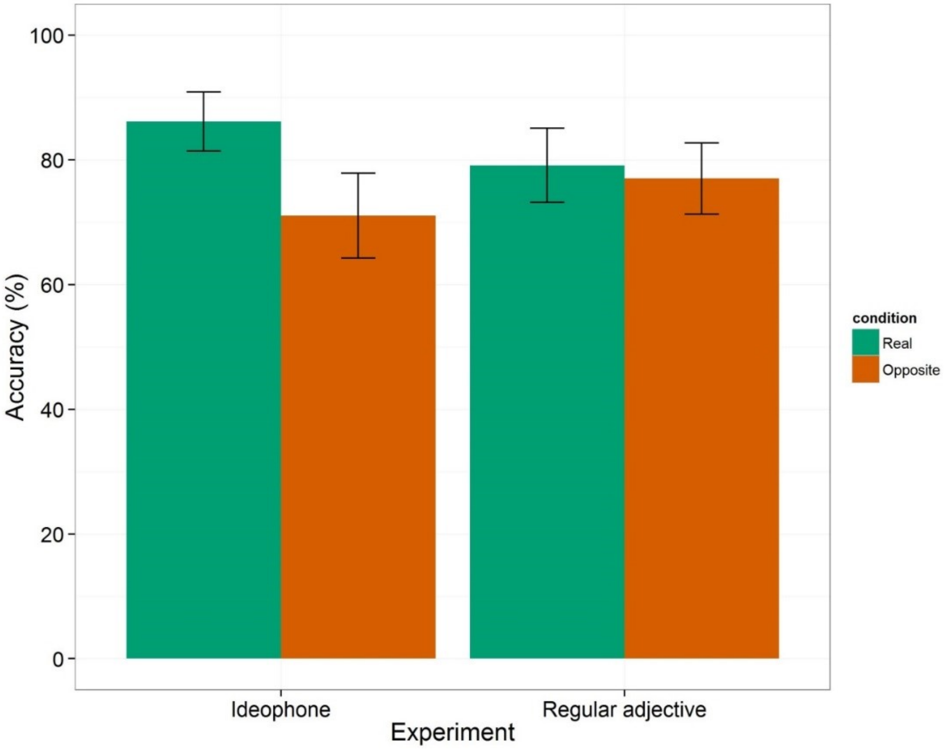


Figure 5: recognition accuracy per condition per experiment with 95% confidence intervals

Discussion

In Experiment 1, we taught Japanese ideophones to Dutch participants with their real and opposite translations, and we found that participants learned the sound-symbolically congruent word pairs (i.e. the ideophone and its real translation) better than the sound-symbolically incongruent word pairs (i.e. the ideophone and its opposite translation). This was corroborated by reaction times, which showed that participants responded faster to the sound-symbolically congruent word pairs. We also found that, despite learning 50% incorrect mappings in the learning task, participants were still able to categorise the ideophones at above chance accuracy in a two-alternative forced choice test afterwards.

In Experiment 2, another set of Dutch participants learned regular Japanese adjectives with their real and opposite translations. Here there was no learning effect at all, nor a difference in reaction times, although participants were still able to categorise the regular adjectives at above chance accuracy in a typical 2AFC test afterwards. These findings show that sound symbolism in Japanese is robustly recognisable by Dutch speakers outside a forced choice paradigm, and that it can be exploited to facilitate word learning. This provides solid empirical grounding for the developmental literature about sound-symbolic bootstrapping, which has tended to use pseudowords rather than real sound-symbolic words.

These findings go beyond previous behavioural work on sound symbolism in two key ways. First, the stimuli more accurately reflect the nature of sound symbolism in natural language, as we use existing sound-symbolic words from a natural language as opposed to deliberately contrastive pseudowords. Second, the task speaks more directly to theories about the role of sound symbolism in learning, as we use a word learning task where participants are free to learn the ideophones and Dutch translations, rather than using a 2AFC word guessing task which limits and shapes the cross-modal associations that participants may form.

The effect of sound symbolism in Experiment 1 is strong and consistent: while some ideophones were answered correctly more often than others, model comparison showed that condition predicted the learning effect when controlling for random effects of ideophones. All but two ideophones were mistaken at least once in the recognition task, and only one ideophone was mistaken in the recognition task by more than half the participants (and even then, only by 17 out of 30); the rest are evenly distributed across those two points. This suggests that the sound-symbolic effect is present across all the ideophones used, affirming the sound-symbolic potential of ideophones.

That we obtained these results despite the open nature of the stimuli and task shows that naturally attested forms of sound symbolism are robust beyond the classic 2AFC paradigm. Of previous studies, only Nygaard et al. (2009) use similar materials and methods. They found judgement accuracy and reaction time differences between words learned with their real translations and with random translations, with participants responding to words learned with their real translations more accurately and more quickly. But, they found no difference in either judgement accuracy or reaction time between words learned with their real and opposite translations. Nygaard et al. argue that this shows that sound-symbolically congruent mappings between form and meaning facilitate word learning. Our Experiment 2 lends support to this interpretation by replicating their finding: for regular adjectives, we find no difference in judgement accuracy and reaction times between real and opposite conditions. Experiment 1, meanwhile, allows us to further explore the nature of sound-symbolic congruence: there, we find large differences in both accuracy judgements and reaction times between real and opposite conditions. We expect that this difference is due to our use of overtly sound-symbolic ideophones, as the sensory sound symbolism in ideophones makes them more transparently iconic than the technically arbitrary and covertly sound-symbolic adjectives which were used in Experiment 2 and in Nygaard et al. (2009).

Comparing Experiment 1 and Experiment 2 (with the caveat that these were done by different sets of participants) suggests that the effect may be driven by both a sound/meaning match providing a mapping boost and a sound/meaning mismatch creating a mapping difficulty. For ideophones (Experiment 1), the difference between the real and opposite conditions is maximal: 86.1% versus 71.1% correct responses. For adjectives (Experiment 2), the difference between real and opposite conditions is minimal, indeed non-significant: 79.1% versus 77% correct responses. That ideophones outperform adjectives in the real condition suggests that the sound/meaning match may provide a mapping boost which helps participants remember the real words, a finding that is in line with the developmental literature on the role of sound symbolism in learning (Imai & Kita, 2014). When this sound/meaning match is not present, participants may default to assuming word arbitrariness, which also works but not quite as well (as seen in the adjectives). That ideophones lead to worse performance than adjectives in the opposite condition suggests that the sound-meaning mismatch may create a mapping difficulty, the converse of the putative mapping boost seen in the real condition. However, it is important to stress that the two experiments featured different groups of participants learning different test items, and the hypothesis of a sound-meaning match mapping boost and a sound-meaning mismatch mapping difficulty can only be tested with the same participants learning both sets of words[2]. More research is needed to uncover the mechanism by which naïve participants come to have different expectations about ideophones versus adjectives in a language they do not speak, but the answer lies likely in a combination of the special morphophonological shapes of ideophones and their relatively specific meanings (Akita, 2011; Dingemanse, 2012).

It is possible that having the 2AFC task after the learning round could bias the participants towards just selecting the words they had learned and remembered, rather than assessing their sensitivity to sound symbolism in general. However, participants could guess the real meanings of the words that they learned in both conditions at the same accuracy. This was the case in both Experiment 1 (75.1% guessing accuracy for ideophones previously learned in the real condition, 69.5% guessing accuracy for ideophones previously learned in the opposite condition, $t=1.31$, $p=0.2$) and Experiment 2 (65.1% guessing accuracy for regular adjectives previously learned in the real condition, 60.8% guessing accuracy for regular adjectives previously learned in the opposite condition, $t=1.12$, $p=0.27$). This suggests that a general sensitivity to sound symbolism persists throughout, and despite, learning opposite mappings. However, it cannot be excluded that this effect may disappear with familiarity with the words, meaning that additional learning rounds and test rounds may bias participants towards selecting answers in the 2AFC task based on what they had learned rather than on their intuition.

It is interesting that the regular Japanese adjectives were also guessed at above chance level in the stimuli selection pre-test (at 55.3% accuracy, compared to 63.1% accuracy in the ideophone stimuli selection pre-test). This result is probably driven by a certain amount of low-level sound symbolism in the mostly arbitrary words, and

---

[2] For the record: the difference between the two match (i.e. real translation) conditions is $t = 1.8$, $p = 0.076$, and the difference between the two mismatch (i.e. opposite translation) conditions is $t = 1.3$, $p = 0.19$. However, we provide these cross-experiment statistics only as a reference point for performing future within-subjects experiments.

the residual levels of informative prosody in the native speaker's recordings; i.e. the kind of covert sound symbolism which is also present in Nygaard et al. (2009). It is probably too simplistic to think of sound symbolism as a binary feature that words or word classes do or don't have; instead it is more useful to think about the *degree* of iconicity (or sound-symbolic congruency) in form-meaning correspondences (Perry, Perlman, & Lupyan, 2015). Here, we have used ideophones as a word class with a relatively high degree of iconicity to study learning effects of sound symbolism, and we have used regular adjectives as a word class with relatively lower degree of iconicity as a control condition to make sure the learning effects really are due to sound symbolism.

Our finding that Dutch participants learn sound-symbolic words with their real meanings better than sound-symbolic words with their opposite meanings raises the questions of how exactly this works, and how universal this is. Future research is required into ideophones from other languages than Japanese with participants with native languages other than Dutch.

Conclusion

This study has shown that Dutch speakers are sensitive to the meanings of Japanese ideophones in both a 2AFC task and a learning task. Sound symbolism appears to provide a mapping boost: when sound and meaning are congruent, learning the link between them is easier. A second experiment with regular adjectives found no such learning effect. This shows that the word classes of natural language may differ in the degree to which they show sound symbolism, with ideophones being more strongly sound-symbolic than regular adjectives. Our results suggest that sound symbolism in ideophones is universally perceivable to at least some extent, and that not only children but also adults can use sound-symbolic cues to bootstrap word learning.

References

Akita, K. (2009). A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics. Kobe University.

Akita, K. (2011). Toward a Phonosemantic Definition of Iconic Words. In P. Michelucci, O. Fischer, & C. Ljungberg (Eds.), Iconicity in Language and Literature 10 (pp. 3–18). Amsterdam: John Benjamins Publishing.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. Cortex, 63, 196–205. http://doi.org/10.1016/j.cortex.2014.08.025

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. Memory & Cognition, 40(1), 83–92. http://doi.org/10.3758/s13421-011-0139-3

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. Cognition, 126(2), 165–172. http://doi.org/10.1016/j.cognition.2012.09.007

Davis, R. (1961). The Fitness of Names to Drawings. a Cross-Cultural Study in Tanganyika. British Journal of Psychology, 52(3), 259–268. http://doi.org/10.1111/j.2044-8295.1961.tb00788.x

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. Language and Linguistics Compass, 6(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. Trends in Cognitive Sciences, 19(10), 603–615. http://doi.org/10.1016/j.tics.2015.07.013

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound symbolism is disrupted in dyslexia: implications for the role of cross-modal abstraction processes. In Proceedings of the 37th Annual Meeting of the Cognitive Science Society.

Hamano, S. (1998). The Sound-Symbolic System of Japanese. Center for the Study of Language and Information ; Tokyo.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33(2-3), 61–83. http://doi.org/10.1017/S0140525X0999152X

Hinton, L., Nichols, J., & Ohala, J. J. (1994). Sound Symbolism. Cambridge: Cambridge University Press.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130298. http://doi.org/10.1098/rstb.2013.0298

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. Cognition, 109(1), 54–65. http://doi.org/10.1016/j.cognition.2008.07.015

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007a). How does it hurt, kiri-kiri or siku-siku? Japanese mimetic words of pain perceived by Japanese speakers and English speakers. In M. Minami (Ed.), Applying theory and research to learning Japanese as a foreign language (pp. 2–19). Newcastle: Cambridge Scholars Publishing. Retrieved from http://eprints.soas.ac.uk/6438/

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007b). What do English speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. Japanese-Language Education around the Globe, 17, 53–78.

Köhler, W. (1947). Gestalt psychology; an introduction to new concepts in modern psychology (Rev. ed.). Oxford, England: Liveright.

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. Journal of Verbal Learning and Verbal Behavior, 10(4), 427–429. http://doi.org/10.1016/S0022-5371(71)80042-7

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. First Language, 34(5), 387–405. http://doi.org/10.1177/0142723714550110

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioural, developmental, and neuroimaging research into sound symbolism. Language Sciences, 6, 1246. http://doi.org/10.3389/fpsyg.2015.01246

Majid, A., & Levinson, S. C. (2010). WEIRD languages have misled us too. Behavioral and Brain Sciences.

Newman, S. S. (1933). Further Experiments in Phonetic Symbolism. The American Journal of Psychology, 45(1), 53. http://doi.org/10.2307/1414186

Nielsen, A., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 67(2), 153–163. http://doi.org/10.1037/a0030553

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. Cognition, 112(1), 181–186. http://doi.org/10.1016/j.cognition.2009.04.001

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. Cognitive Science, 33(1), 127–146. http://doi.org/10.1111/j.1551-6709.2008.01007.x

Oak, S. (1996). Pikachu. In Pokédex. Nintendo.

Occelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013). The takete–maluma phenomenon in autism spectrum disorders. Perception, 42(2), 233 – 241. http://doi.org/10.1068/p7357

Oda, H. (2000). An embodied semantic mechanism for mimetic words in Japanese. Indiana University, Indiana.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. Frontiers in Psychology, 1(E227). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3153832/

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition. PLoS ONE, 10(9), e0137147. http://doi.org/10.1371/journal.pone.0137147

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia -- A window into perception, thought and language. Journal of Consciousness Studies, 8(12), 3–34.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. Brain and Language, 128(1), 18–24. http://doi.org/10.1016/j.bandl.2013.11.002

Sapir, E. (1929). A study in phonetic symbolism. Journal of Experimental Psychology, 12(3), 225–239. http://doi.org/10.1037/h0070931

Shogakukan. (2001). Shogakukan Progressive English-Japanese Dictionary (3rd edition). Shogakukan.

Thompson, P., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. The Quarterly Journal of Experimental Psychology, 64(12), 2392–2404. http://doi.org/10.1080/17470218.2011.605898

Vigliocco, G., & Kita, S. (2006). Language-specific properties of the lexicon: Implications for learning and processing. Language and Cognitive Processes, 21(7-8), 790–816.

Voeltz, F. K. E., & Kilian-Hatz, C. (Eds.). (2001). Ideophones. Amsterdam: John Benjamins.

Yoshida, H. (2012). A Cross-Linguistic Study of Sound Symbolism in Children’s Verb Learning. Journal of Cognition and Development, 13(2), 232–265. http://doi.org/10.1080/15248372.2011.573515

Chapter 5

How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping.

Published as:
Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How Iconicity Helps People Learn New Words: Neural Correlates and Individual Differences in Sound-Symbolic Bootstrapping. *Collabra*, 2(1).

Abstract

Sound symbolism is increasingly understood as involving iconicity, or perceptual analogies and cross-modal correspondences between form and meaning, but the search for its functional and neural correlates is ongoing. Here we study how people learn sound-symbolic words, using behavioural, electrophysiological and individual difference measures. Dutch participants learned Japanese ideophones —lexical sound-symbolic words— with a translation of either the real meaning (in which form and meaning show cross-modal correspondences) or the opposite meaning (in which form and meaning show cross-modal clashes). Participants were significantly better at identifying the words they learned in the real condition, correctly remembering the real word pairing 86.7% of the time, but the opposite word pairing only 71.3% of the time. Analysing event-related potentials (ERPs) during the test round showed that ideophones in the real condition elicited a greater P3 component and late positive complex than ideophones in the opposite condition. In a subsequent forced choice task, participants were asked to guess the real translation from two alternatives. They did this with 73.0% accuracy, well above chance level even for words they had encountered in the opposite condition, showing that people are generally sensitive to the sound-symbolic cues in ideophones. Individual difference measures showed that the ERP effect in the test round of the learning task was greater for participants who were more sensitive to sound symbolism in the forced choice task. The main driver of the difference was a lower amplitude of the P3 component in response to ideophones in the opposite condition, suggesting that people who are more sensitive to sound symbolism may have more difficulty to suppress conflicting cross-modal information. The findings provide new evidence that cross-modal correspondences between sound and meaning facilitate word learning, while cross-modal clashes make word learning harder, especially for people who are more sensitive to sound symbolism.

Introduction

Iconicity, or the resemblance-based mapping between aspects of form and meaning, has long been marginalised in linguistic research due to the predominance of arbitrariness, where there is no connection between the form of a word and aspects of its meaning other than social convention (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015). For example, there is nothing iconic about the arbitrary word *dog*; *d* does not mean "four-legged", *o* does not mean "pet", and *g* does not mean "likes rolling in muddy puddles". Nothing about the form of the word represents the real world meaning. In contrast, the Siwu words *pimbilii* ('small belly') and *pumbuluu* ('enormous round belly') use the vowel space to iconically depict the size of the referent's belly. In British Sign Language, the sign for 'tree' is also iconic, as it features the primary forearm raised, representing the trunk, with the hand open and the fingers spread, representing the branches and leaves (Perniss & Vigliocco, 2014).

Recent psychological research has shown that iconicity plays a bigger role in language than traditionally thought (Lockwood & Dingemanse, 2015; Perniss, Thompson, & Vigliocco, 2010), and that people are sensitive to sound symbolism (which is iconicity specifically for spoken languages) in psycholinguistic tasks. However, much of this research is based on two-alternative forced choice paradigms with pseudowords which are deliberately constructed to maximise iconic contrasts. For a more detailed picture of how sound symbolism works in natural language, pseudoword experiments will need to be supplemented with work using real sound-symbolic words. In spoken languages, one source of sound-symbolic words is the lexical class of ideophones, which are marked words which depict sensory imagery (Dingemanse, 2012; Voeltz & Kilian-Hatz, 2001).

Behavioural experiments with ideophones have mostly tended to show that people can guess the meanings of ideophones at above chance levels, and that this is modulated by articulation and prosody. Oda (2000) showed that English speakers could guess at above chance levels which Japanese ideophones and English translations went together when hearing the words pronounced by a Japanese speaker, and that their accuracy improved when they articulated the ideophones themselves. Iwasaki et al. also showed that English speakers were sensitive to the meanings of Japanese ideophones, and that English speakers' judgements on the semantic dimensions of the event depicted by an ideophone were broadly consistent with those of Japanese speakers. Meanwhile, Kunihira (1971) found that English speakers could guess the meanings of apparently arbitrary Japanese words better than chance when hearing the words in a monotone voice, and better still when hearing the words in an expressive voice. This shows that even arbitrary words may have residual levels of sound symbolism in them, and that prosody is an important factor in the perception sound symbolism.

Nygaard et al. (2009) followed up Kunihira's study with a word learning experiment. They found that English participants were faster and more accurate at remembering Japanese words taught with their actual English translation than a random English translation, and that there was no difference when learning words with their opposite translations. Nygaard et al. argue that "the sound structure of spoken language may engage cross-modal perceptual-motor correspondences that permeate the form, structure, and meaning of linguistic communication" and that the learners in their experiment were unconsciously able to "exploit non-arbitrary relationships in the service of word learning and retrieval". While we are sympathetic to Nygaard et al.'s arguments, a limitation of their study is that using a variety of nouns, adjectives, and verbs with a variety of prosodic contours and morphophonological structures obscures the many potential sources of sound symbolism that their participants may have identified. Moreover, some of the stimuli used lend themselves more obviously to real and opposite conditions (such as using "slow" as the opposite of *hayai*, which means "fast") than others (such as using "gold" as the opposite of *tetsu*, which means "iron"). Finally, while the words that Kunihira and Nygaard et al. used in their experiments (and many words in many languages in general) may contain a certain degree of sound-symbolic mappings, they are generally considered to be arbitrary.

In Lockwood, Dingemanse, and Hagoort (2016), we ran a learning experiment similar to the Kunihira and Nygaard studies, but strictly with Japanese ideophones, which were controlled for length, grammatical category, and morphophonological structure, and which are strongly sound-symbolic. We showed that Dutch adults learned novel Japanese ideophones better when they were learned with their real Dutch translations (i.e. when there was a sound-symbolic relationship between form and meaning) than when they were learned with their opposite Dutch translations (i.e. when there was either no match or a mismatch between form and meaning). We then informed participants of the manipulation, and asked them to choose what they thought the best translation would be in a two-alternative forced choice task. Despite the learning task, the participants were still sensitive to the ideophones' meanings and guessed well above chance at 72% accuracy. Meanwhile, we ran the same manipulation with a set of arbitrary adjectives —i.e. adjectives which are not ideophones and are not considered sound-symbolic— with a second group of participants. Participants were able to guess the meanings of the words in a two-alternative forced choice test above chance at 63% accuracy, but the learning effect disappeared completely; participants remembered the adjectives with their real translations at the same level of accuracy as the adjectives with their opposite translations, echoing Nygaard et al.'s (2009) findings. We used Japanese for consistency with earlier studies and because Japanese is probably the most well-documented language with an extensive set of ideophones (Akita, 2009; Hamano, 1998), but given the typological unity of ideophones (Dingemanse, 2012; Samarin, 1965) we are fairly confident that this effect would hold in any language with ideophones.

There has been relatively little neuroimaging work on sound symbolism involving real words. In a sentence reading experiment with native Japanese speakers, Lockwood & Tuomainen (2015) found that ideophones elicit a greater P2 component and elicit a larger late positive complex (LPC) compared to arbitrary words. They argue that the P2 reflects the multisensory integration of sounds and the associated sensory representations, and that the LPC may reflect higher processing demands of ideophones. In fMRI experiments, Revill et al. (2014) and Kanero et al. (2014) have both found that sound-symbolic words activate certain brain areas more strongly than non-sound-symbolic words. Revill et al. used words from a variety of languages (some of them historically related) and from a variety of word classes, and labelled the words which English speakers were better able to guess the meanings of as "sound-symbolic", while words that they guessed at chance were labelled as "non-sound-symbolic". The sound-symbolic words elicited more activation than the non-sound-symbolic words in intraparietal areas associated with cross-modal and synaesthetic processing. Kanero et al. compared Japanese ideophones with arbitrary words when participants viewed matching or mismatching videos of motions and images of shapes. They found that the ideophones uniquely activated the right posterior superior temporal sulcus, and that this activation was greater when the ideophones and the videos/images were rated as better matching. They speculate that the right posterior STS integrates the processing of linguistic and environmental sounds. There is very little event-related potential (ERP) research on sound symbolism in real words, and Kanero et al. and Revill et al. make similar arguments about different brain areas in fMRI research. This means that more neuroimaging work is needed in order to work out how the brain processes sound symbolism.

Here we build on this work, extending it in two ways to advance our understanding of sound symbolism. First, we used ideophones, words considered strongly sound-symbolic or iconic by both linguists (Diffloth, 1972; Dingemanse, 2012) and native speakers (Kita, 1997; Kunene, 1965). Using a more unified and linguistically and prosodically homogeneous set of words makes it easier to eliminate possible confounds and be confident that any effect we find is a reliable indicator of sound symbolism. We repeated the behavioural task in Lockwood et al. (2016) with minor alterations to measure the participants' brain activity using EEG (electroencephalography). As Lockwood et al. showed that there was no learning effect with regular arbitrary adjectives, only ideophones were used in the current study.

Second, we analysed event-related potentials (ERPs) to explore the neural mechanisms underlying the processing of sound-symbolic words. We used ERPs to look at the time course of the neural effect; if an early effect was present, as in Kovic et al.'s (2010) study with pseudowords, this would suggest that the effect is based on differences in the processing of the sensory properties of the stimuli, whereas if the effect was much later, it would suggest a more linguistic mechanism. It is possible that there are both sensory and linguistic effects, as suggested in ERP experiments by Lockwood and Tuomainen (2015) and Sučević et al. (2015).

Coupling behavioural data and brain imaging allows us to investigate possible individual differences in sound-symbolic sensitivity. The topic of individual differences (Levinson, 2012) has barely been broached in the sound-symbolism literature so far, but is likely to be of key importance in the quest for causal models of sound-symbolism.

We hypothesised that we would behaviourally replicate Lockwood et al. (2016), namely that participants would learn the ideophones in the real condition better than in the opposite condition and that participants would still be sensitive to the meanings of ideophones in the two-alternative forced choice task afterwards despite the learning rounds. We also predicted that there would be a correlation between the reaction time and accuracy of judgement of ideophones, in that the more accurately guessed ideophones would also be more quickly guessed. As for the ERP results, since the few sound symbolism ERP studies so far have found different components, we used a non-parametric cluster-based permutation test to investigate the data before analysing particular windows. Finally, we investigated individual differences in the data by looking at the relation between the ERP effect size, the memory/learning performance of the task, and behavioural measures of sensitivity to sound symbolism per participant. We did this in order to see whether the effect was more related to participants' sensitivity to sound symbolism or more related to participants' general task performance.

Methods

Stimuli

This experiment used the same paradigm as Lockwood et al. (2016).

We used 38 Japanese ideophones with a reduplicated CVCV-CVCV pattern. Ideophones and Dutch translations were matched for word length and characters in common across conditions. Dutch translations across conditions were additionally matched for word frequency. Participants learned the correct translations to 19 ideophones and the opposite translations to the other 19 ideophones. In a published pretest using a fully counterbalanced set of stimuli (Lockwood et al. 2016), we found a main effect of real vs. opposite condition in both groups. As counterbalancing made no difference to the results, the stimuli we use here are consistent across participants; all learned *fuwafuwa* as *pluizig*, for example.

REAL condition		OPPOSITE condition	
ideophone	translation	ideophone	translation
fuwafuwa ("fluffy")	pluizig ("fluffy")	kibikibi ("energetic")	futloos ("tame, tired")
boroboro ("worn out")	versleten ("worn out")	ukiuki ("happy")	verdrietig ("sad")

Table 1: example stimuli for each condition

Procedure

Participants were told that they were going to learn 38 Japanese words, and that they had to remember the word pairs for a recognition test straight after the learning rounds.

After the test, participants were informed that half the words they had learned were correct, but half were the opposite meaning. We then asked them to ignore what they had just learned and instead choose which translation they felt was more natural for each ideophone during the 2AFC task.

Participants saw each ideophone and translation once in a learning round; there were two learning rounds in total. The order of Dutch words and ideophones was randomised for each round and for each participant. We used Presentation to present stimuli and record responses.

The initial Dutch word was presented for 1000ms with 100ms of jitter each way (i.e. between 900ms and 1100ms), followed by a fixation cross for 1000ms with 100ms of jitter. As the ideophone was played over the speakers, a blank screen was presented for 2000ms with 200ms of jitter. This was again followed by a fixation cross. The final screen with the ideophone and its Dutch meaning was presented until participants were happy to move onto the next item. Between trials, a blank screen was presented for 1000ms with 200ms of jitter, followed by a fixation cross for 1000ms with 100ms of jitter to announce the beginning of the next trial.

When it came to the test round, participants were presented with either the word pairs that they had learned (for example, *fuwafuwa* and *pluizig* in the real condition, and *kibikibi* and *futloos* in the opposite condition), or a pseudo-randomised pairing of ideophones and translations which they had seen before. These pairings were pseudo-randomised to ensure that the meanings were semantically unrelated (for example, the Japanese *fuwafuwa*, learned as "fluffy", and the Dutch *kortaf*, meaning "curt"). Participants were instructed to indicate whether this was a word pair they had learned by answering Yes (left CTRL key) or No (right CTRL key). Pairs requiring a Yes response made up 50% of the trials. As in the learning round, participants saw the Dutch word first, then heard the Japanese ideophone for 2000ms. Then, instead of seeing a fixation cross, they saw a question mark. Participants were asked to respond as soon as possible after seeing the question mark.

Timings in the test stage were identical to the learning stage. The question mark was displayed until participants responded, at which point a blank screen was presented, followed by a fixation cross to announce the beginning of the next trial. In order to ensure enough trials for ERP analysis, the test stage was twice as long as in Lockwood et al. (2016), so that there were 38 trials per condition (i.e. 19 ideophones with their real translation, 19 ideophones with their opposite translation, and 38 ideophones with a pseudo-randomised wrong translation, all repeated).

After the test round, we implemented a two-alternative forced choice task as a separate measure of sound-symbolic sensitivity. This was to see if, despite the learning phase, participants were still able to make decisions based on the sound symbolism of the ideophones. Participants heard the ideophone, and then saw the two possible Dutch translations; they selected the translation by pressing the left CTRL key for the translation on the left and the right CTRL key for the translation on the right. Timings were identical to the learning and test stages.

The full experiment is illustrated in Figure 1.

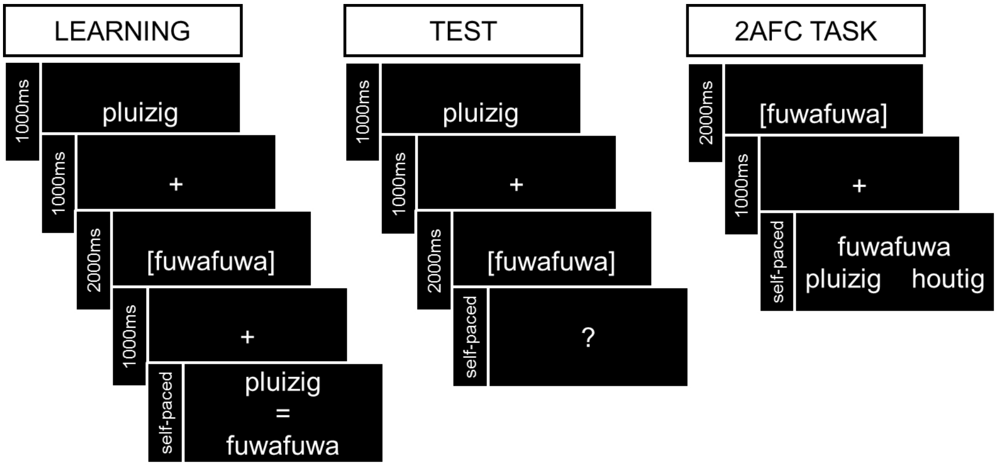


Figure 1: Learning, test, and 2AFC procedure

Participants

We tested 40 native Dutch speaking participants (10m, 30f) aged 18-29 (mean: 21y 7m) with normal or corrected-to-normal vision, recruited from the MPI participant database. All participants had no knowledge of Japanese, and were students at either the Radboud University or the Hogeschool van Arnhem en Nijmegen. Participants gave informed written consent to take part in the experiment. The experiment was approved by the Ethics Committee for Behavioural Research of the Social Sciences Faculty at Radboud University Nijmegen in compliance with the Declaration of Helsinki. Participants and were paid 8 Euro per hour for their participation.

In order to make sure that we were testing ERPs from participants who had learned the words, we discarded five participants who scored under 60% in the test round and could have just been guessing the answers. A further six participants were discarded due to excessive artefacts (affecting more than 25% of trials). This left 29 participants in the final dataset (7m, 22f; 19-28 years old, mean 21y9m; 24 right-handed, 5 left-handed).

EEG recording

EEG was recorded from 61 active Ag/AgCl electrodes, of which 59 were mounted in a cap (actiCap), referenced to the left mastoid. Two separate electrodes were placed at the left and right mastoids. Blinks were monitored through an electrode on the infraorbital ridge below the left eye. The ground electrode was placed on the forehead. Electrode impedance was kept below 10 kΩ. EEG and EOG recordings were amplified through BrainAmp DC amplifiers with a bandpass filter of 0.016–100 Hz, digitised on-line with a sampling frequency of 500 Hz, and stored for off-line analysis.

ERP analysis

Automatic artefact rejection in BrainVision Analyzer discarded all segments with activity exceeding ±75 µV. In six of the 29 participants used for the ERP analysis, between one and four individual electrodes were removed and interpolated due to faulty connections. ERPs were timelocked to the onset of the ideophone recording. Across the 29 participants used for all analyses reported in this paper, 13.1% of ideophone trials were rejected due to artefacts.

As previous sound symbolism studies using ERPs have found mixed results, we used a non-parametric cluster-based permutation test in Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). This investigated the entire epoch to establish whether there was a difference between conditions at any given point while correcting for multiple comparisons, and highlighted time windows of interest to analyse. We then ran ANOVAs on mean amplitudes in individual time windows of interest.

Results

Behavioural Results

Main experiment

Participants made more recognition mistakes in the opposite condition than in the real condition; participants correctly remembered the real word pairing 86.7% of the time (95% CIs: 82.92% - 90.41%), but correctly remembered the opposite word pairing only 71.3% of the time (95% CIs: 65.19% - 77.46%). This is shown in Figure 2 below, presented in this way rather than as a histogram with error bars in order to better represent the spread of data (Weissgerber, Milic, Winham, & Garovic, 2015). Only six out of 29 participants did not show an advantage for the real condition over the opposite condition. Four participants scored higher in the opposite condition than the real condition (with a mean difference of 3.95 percentage points), and two participants had equal scores in both conditions.

As the dependent variable was binary —correct or incorrect— we analysed the responses using a mixed-effects logit model with the glmer function of the lme4 (versions 1.1-8) package in R. The data was modelled by including a per-participant and per-ideophone random adjustment to the fixed intercept with a random slope for the fixed effect by participant. The condition was sum contrast coded.

Model comparison showed a random effect by ideophone (log likelihood difference = 21.3, χ2 = 42.64, df = 1, p <0.001). That means that some ideophones were answered correctly more often than others. However, even when controlling for this random effect by ideophone, model comparison showed a significant fixed effect of condition (β = −0.5514, log likelihood difference = 8.2, χ2 = 16.44, df = 1, p <0.001). The model estimated that ideophones learned in the real condition were answered 8.1 percentage points more accurately than ideophones learned in the opposite condition.

There were also significant differences in reaction times between conditions, with participants responding faster to ideophones in the real condition (mean RT = 958ms ± 95ms CIs) than the opposite condition (mean RT = 1262ms ± 86ms CIs) (t = -5.00, p <0.001, Cohen's d = -1.63). This difference existed even when only analysing correctly answered trials (t = -4.58, p <0.001, Cohen's d = -1.49), and so is not just a speed/accuracy trade off. There was also a strong correlation between the number of correct responses per ideophone and the speed of the reaction to that ideophone; the better an ideophone was remembered, the faster it was responded to (r = −0.71, p <0.001). However, there was no correlation between the number of correct responses per participant and reaction times (r = -0.11, p = 0.57, Cohen's d = 0.63), meaning that more accurate participants were not necessarily faster at responding.

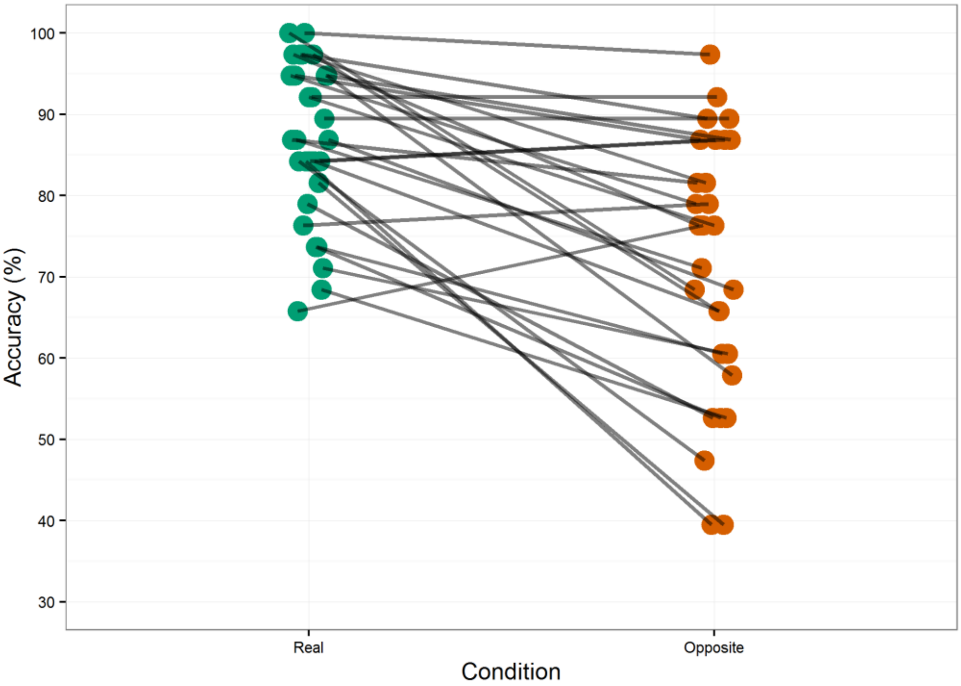


Figure 2: Accuracy per condition. Dots represent a participant's performance in each condition, lines connect the same participant across conditions.

This closely replicates the results from Lockwood et al. (2016), as shown in Figure 3.

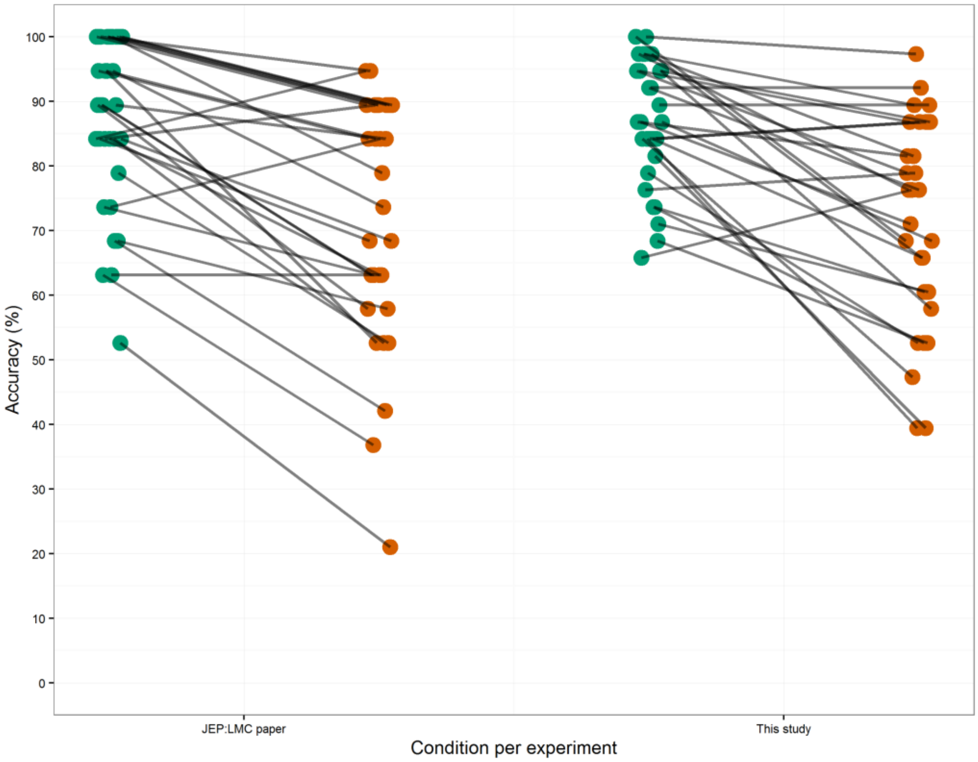


Figure 3: Accuracy per condition, in comparison to our previous behavioural study

Post-experiment sound symbolism sensitivity check

In the sound symbolism sensitivity check after the experiment, participants guessed the real meanings of the Japanese words with 72.96% accuracy, which was comfortably above chance ($\mu = 0.5$, $t = 13.86$, $df = 28$, $p<0.0001$, 95% CIs = 69.56%-76.35%, Cohen's $d = 5.24$). Only one participant guessed the ideophones at 50% accuracy, and 27 out of 29 participants guessed at least 24 out of 38 ideophones correctly. We checked to see if participants who guessed more accurately also guessed faster, but there was no link between reaction times and accuracy ($r=0.07$, $n=29$, $p=0.73$). Only three ideophones were guessed at below 50% accuracy (*hiyahiya* at 41.4%, *morimori* at 44.8%, *gowagowa* at 48.4%). We also checked to see if ideophones which were guessed more accurately were also guessed more quickly. There was a correlation between reaction time and the mean accuracy at which the ideophone was guessed ($r= -0.46$, $n=38$, $p=0.0037$). This is shown in Figure 4.

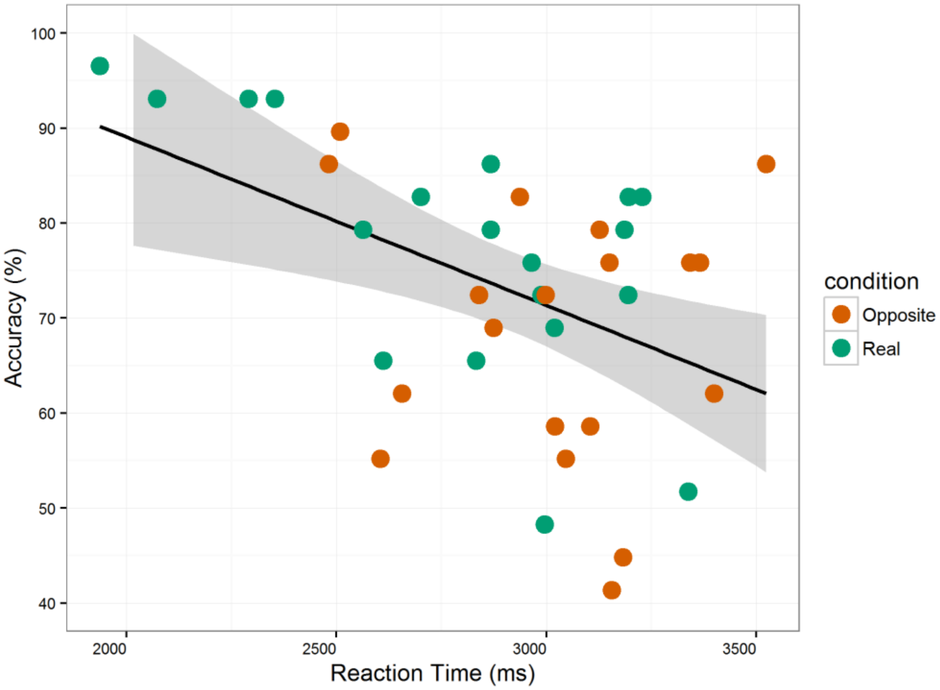


Figure 4: Scatterplot showing correlation between mean accuracy per ideophone and reaction time. Dots represent ideophones, colours represent the condition the ideophones were learned in.

One might ask whether the two-alternative forced choice test is affected by taking place *after* the learning round, as participants might continue to select the words they had learned and maybe only change a few decisions. Participants guessed the real meanings of the words they had previously learned in the real condition at 77.3% (95% CIs: 70.9% - 83.7%), and they guessed the real meanings of the words they had previously learned in the opposite condition at 68.6% (95% CIs: 61.9% - 75.3%), and this is shown in Figure 5. This suggests that participants were still sensitive to sound symbolism, especially as they picked the correct translation of the ideophones originally taught in the opposite condition 68.6% of the time despite being taught explicitly otherwise. However, they may have found it harder to reverse this learning than they did to reëvaluate the ideophones they had learned in the real condition; there was a trend towards guessing ideophones previously learned in the real condition more accurately than ideophones previously learned in the opposite condition ($t = 1.9665$, $p = 0.057$, Cohen's $d = 0.64$). This is in line with our predictions in Lockwood et al. (2016) that further exposure to ideophones and learned translations decreases the ability to reëvaluate sound-symbolic mappings.

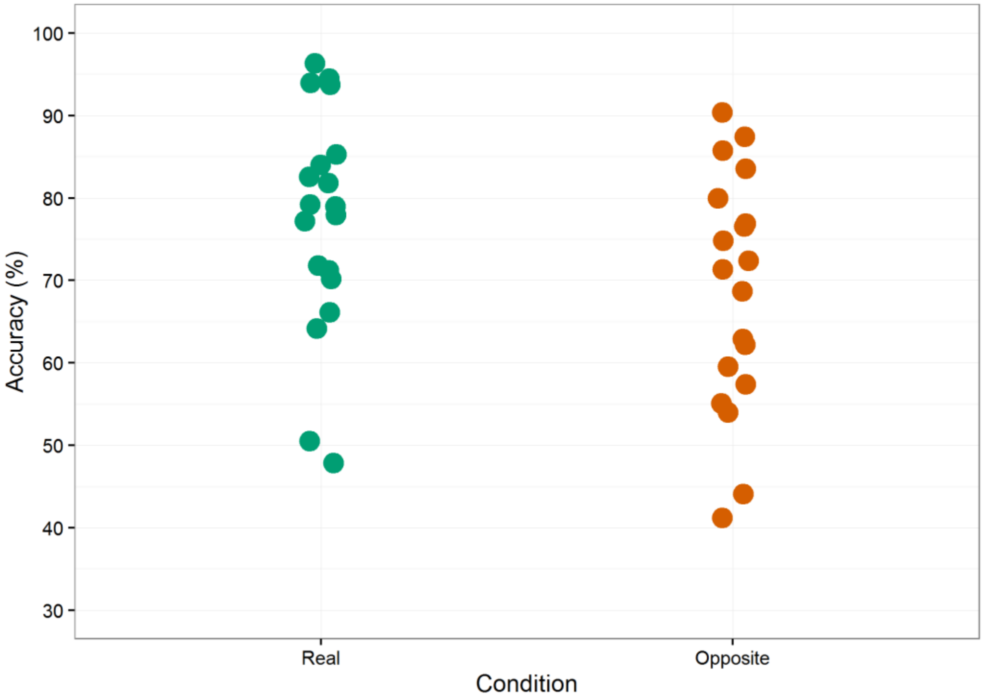


Figure 5: Scatter plot showing the lack of difference in baseline guessing accuracy depending on the condition the ideophone had previously been learned in. Dots represent ideophones.

Behavioural measures of sound symbolism

Finally, we contrasted the two behavioural measures of sound symbolism in the experiment: the two-alternative forced choice task, and the difference in test scores between the real condition and the opposite condition per participant. Participants who are more sensitive to sound symbolism should find it easier to remember the ideophones in the real condition and harder to remember the ideophones in the opposite condition; therefore, participants who scored higher in the two-alternative forced choice task should also have a greater disparity in their test scores between conditions.

The two measures were ranked and showed a Spearman correlation (r = 0.42, p = 0.0251), which suggests that people who are sensitive to sound symbolism when asked to guess a word's meaning are more likely to be affected by that sensitivity during word learning. The correlation is plotted in Figure 6 below.

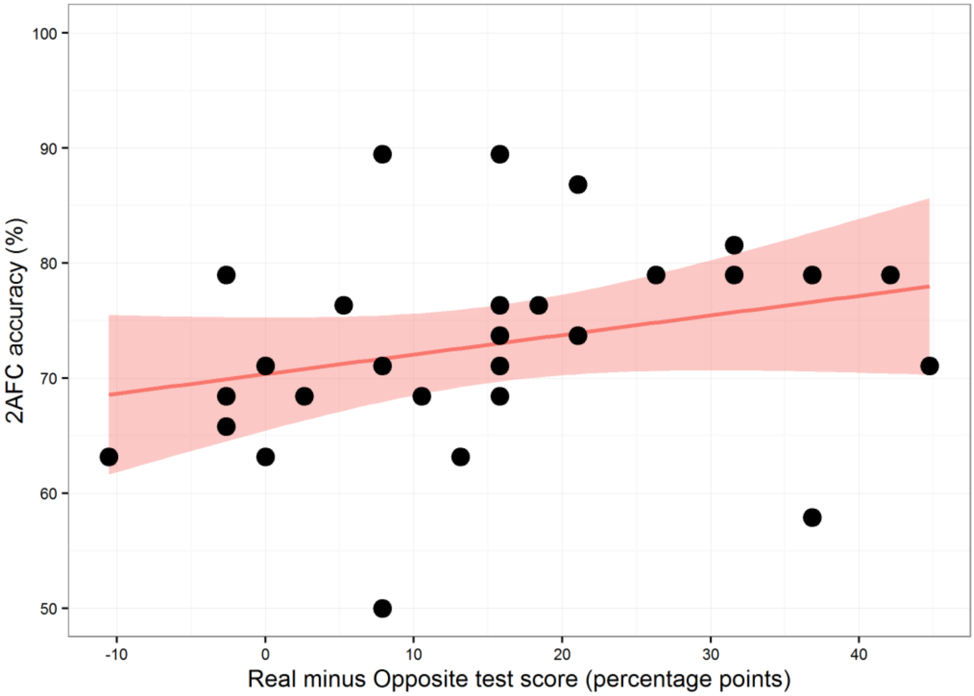


Figure 6: Scatter plot of test score difference and 2AFC task accuracy. Dots represent participants.

ERP results

We examined the ERPs from the participants' passive exposure to the ideophones in the learning rounds and from the participants' exposure to the ideophones during the test round.

Somewhat to our surprise, there was no effect of sound symbolism when participants heard the ideophones during the learning rounds. ERPs were timelocked to the onset of the recording of the ideophone in the learning rounds, but there was no effect when looking at the first learning round, the second learning round, or both together. However, there was a considerable effect in the test round.

In the ERPs from the test rounds, we first ran a cluster-based permutation test with 3000 randomisations in Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) to establish whether there were any differences between real and opposite conditions across the entire averaged epoch. The cluster-based permutation revealed that there was a significant difference between the two conditions, and that this difference was driven by one cluster starting at 320ms and ending at 786ms (p = 0.0027).

Averaged ERP mean amplitudes from nine parietal electrodes (C30, C29, C28, C1, C3, C4, C33, C34, C35) are shown below in Figure 7, and topographic plots of the difference between conditions are shown in Figure 8. The ERPs are time-locked to the onset of the ideophone. Shading around the ERP lines shows 95% confidence intervals. The topographic plots are calculated by subtracting the opposite condition measurements from the real condition measurements.

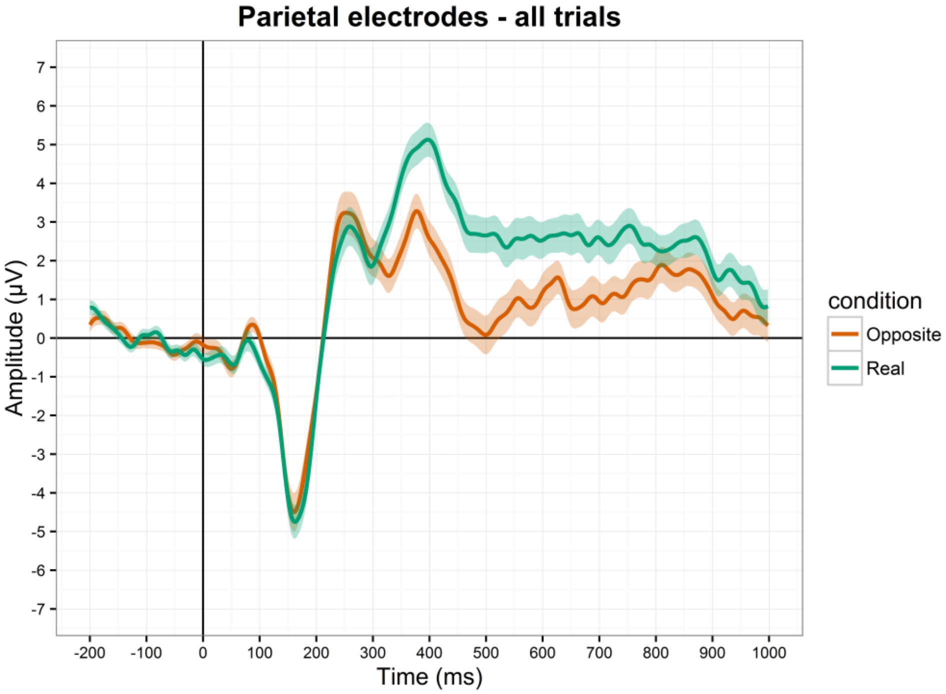


Figure 7: ERPs from all test round trials at the parietal electrodes

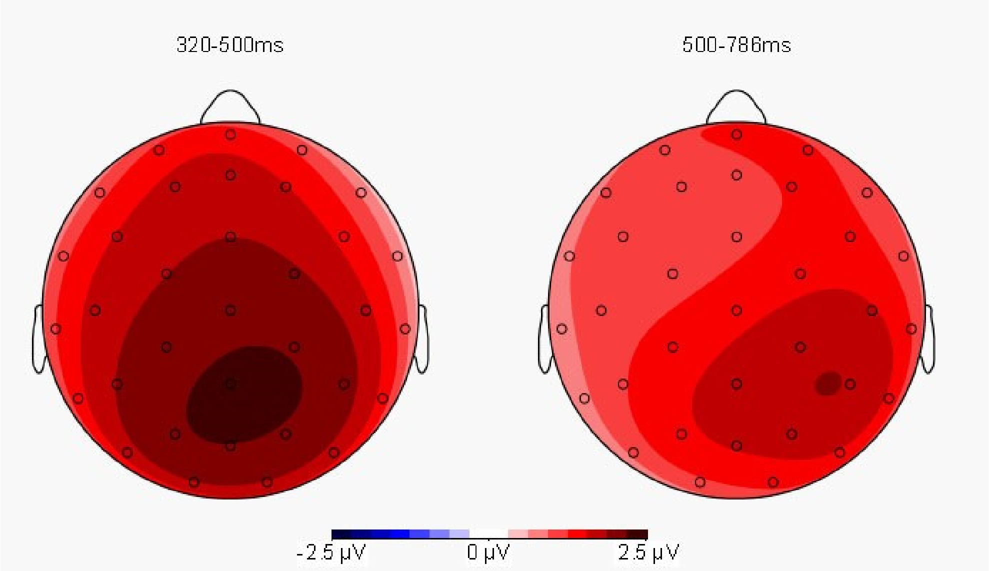


Figure 8: topographic plots of the real minus opposite difference wave in the test round.

We used the cluster and inspection of the waveforms to inform our selection of time windows for further analysis; a P3 effect from 320ms to 500ms, and a late positive complex from 500ms until the end of the cluster at 786ms.

We averaged electrode amplitudes across the midline and four quadrants (left anterior, right anterior, left posterior, right posterior) and ran within-subject 2x5 ANOVAs on the three time windows. In both windows, there was a significant main effect of condition, with ideophones in the real condition eliciting greater a P3 (F = 16.99, df = 1,28, p = 0.0003) and late positive complex (F = 8.96, df = 1,28, p = 0.0057). Interactions between condition and quadrant were not significant for the P3 (p = 0.051) or late positive complex (p = 0.17). Although the interaction was not significant, the P3 effect was greatest in parietal areas in the posterior quadrants.

This analysis included all trials, regardless of whether the participants answered them correctly. To double check, we also analysed only trials which participants answered correctly. Across the 29 participants, an additional 18.6% of trials were rejected due to incorrect responses. Statistical analyses revealed similar results to the analyses of all trials, but all effects were weaker due to having fewer trials.

The effect appears to be centro-parietal according to the topoplots in Figure 8, and therefore it is unlikely that lateralisation of language function due to handedness would make any difference to the data. However, we repeated the analyses when excluding the five left-handed participants in the data to double check. Statistical analyses revealed similar results to the analyses of all participants, but all effects were weaker due to having fewer trials.

Accordingly, all statistics reported in the rest of the paper include all trials and all participants.

These analyses are summarised in Table 2. Here, ges refers to the generalised eta squared measure of effect size.

	All trials			Correct answers only		
Cluster-based permutation test	320-786ms p=0.0027			380-594ms p=0.011		
ANOVA window	F	p	ges	F	p	ges
320-500ms	16.99	0.00030	0.056	7.96	0.0087	0.037
500-786ms	8.96	0.0057	0.032	7.86	0.0091	0.033
320-500ms (LH removed)	15.47	0.00066	0.060	5.84	0.024	0.032
500-786ms (LH removed)	5.12	0.033	0.021	3.30	0.082	0.015

Table 2: table of main effect of condition results

Correlations between behavioural and neurophysiological results

The ERP difference between conditions during the test round could be driven by the sound-symbolic nature of the ideophones, but it could also be an unrelated learning or memory effect. To tease the two apart, we ran individual differences ranked correlations between ERP results and our two behavioural measures: differences in test scores across conditions in the learning task, and accuracy in the sound-symbolic sensitivity task.

If the P3 amplitude in this experiment is related to how easy the ideophones were to learn in the real versus the opposite condition, then the average P3 amplitude per condition per participant should correlate with the participant's test score in that condition in the learning task. However, there was no correlation between P3 amplitude and test score in the real condition or in the opposite condition, which suggests that the ERP effect may be related to something other than ease of learning or recognition.

The P3 amplitude difference between conditions may instead reflect the participants' sensitivity to sound symbolism. If so, then participants who were more sensitive to sound symbolism —as measured in the separate sound-symbolic sensitivity check— should show a greater difference between P3 amplitude peaks than participants who were less sensitive to sound symbolism.

We calculated the P3 effect magnitude by subtracting the average amplitude in the opposite condition from the average amplitude in the real condition per participant. We then correlated the effect magnitudes with participants' two-alternative forced choice accuracy scores from the sound-symbolic sensitivity check. These measures were significantly correlated ($r = 0.42$, $p = 0.0236$), meaning that participants who are better at guessing the meanings of ideophones show a greater P3 effect.

Since the two-alternative forced choice task was significantly correlated with the test score difference between conditions, we also correlated test score differences with P3 amplitude differences across participants. This suggested the same relationship, but was not significant ($r = 0.34$, $p = 0.067$).

Taken together, the correlations between behavioural measures of sound-symbolic sensitivity and P3 amplitude difference between conditions suggests that the P3 effect found in this experiment is related to an individual's sensitivity to sound symbolism. The lack of a relationship between the P3 amplitude and test score per condition goes some way towards ruling out a non-sound-symbolic learning or recognition effect.

To explore this further, below are plotted the same ERPs for participants grouped according to their score in the two-alternative forced choice task. The top half of participants all scored above the mean of 72.96% ($N = 15$), and the mean of their scores was 79.65%. The bottom half of participants all scored below the mean of 72.96% ($N = 14$), and the mean of their scores was 65.79%. Despite the bottom half of participants still scoring comfortably above chance in the sound-symbolic sensitivity task, the P3 effect from the learning task all but disappeared, as shown in Figure 9 and Figure 10.

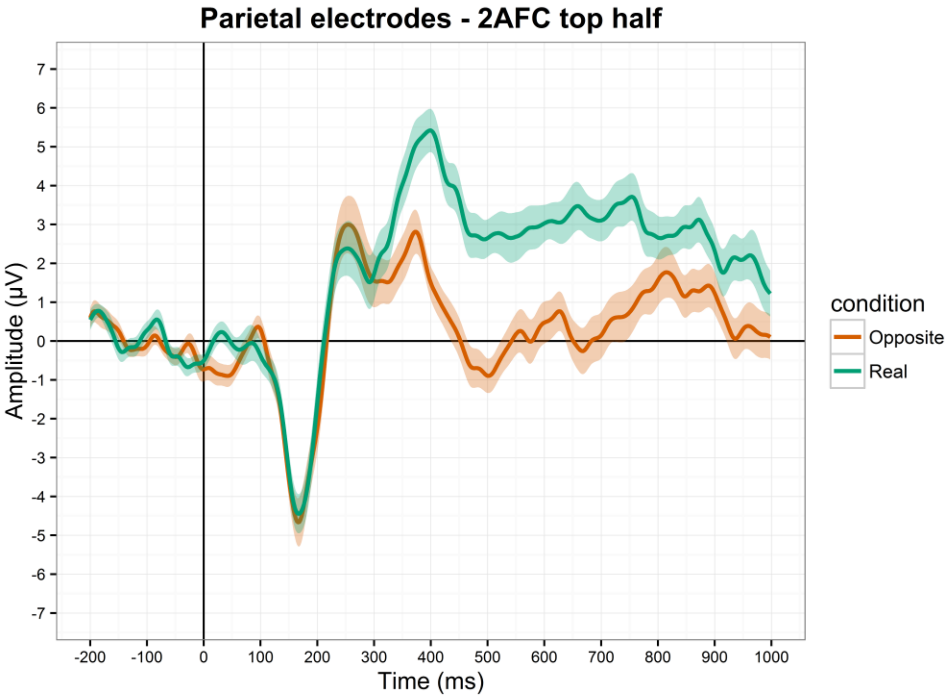


Figure 9: ERPs for the top-performing 15 participants in the 2AFC task measuring sound-symbolic sensitivity

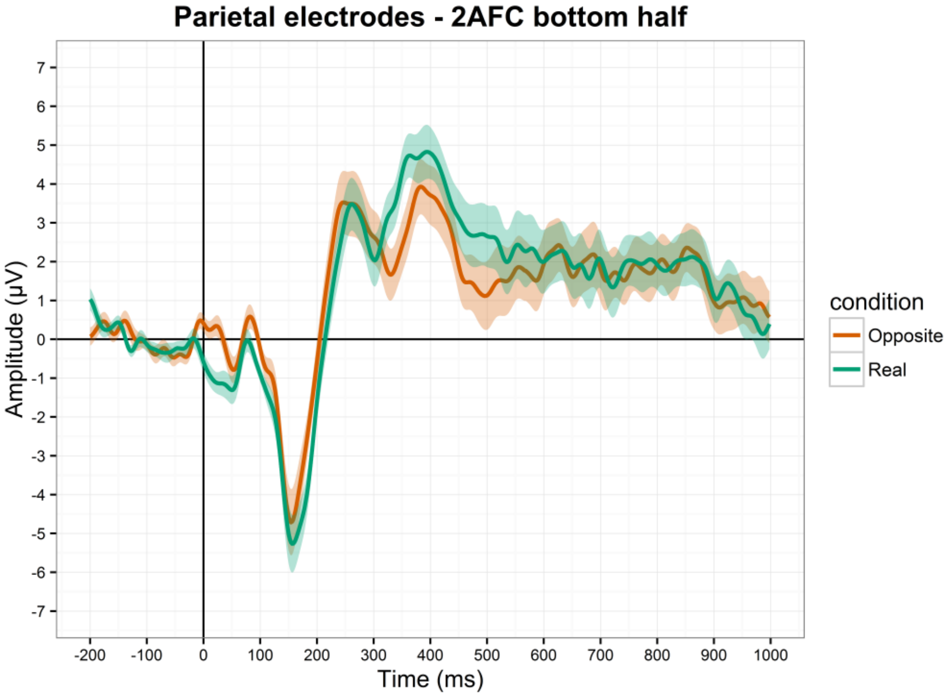


Figure 10: ERPs for the bottom-performing 14 participants in the 2AFC task measuring sound-symbolic sensitivity

When we re-run the ERP ANOVAs for the 2AFC top half and 2AFC bottom half groups separately, the effect is much smaller, indeed not significant, for the 2AFC bottom half group (F = 4.13, p = 0.063, ges = 0.019) while remaining consistent for the 2AFC top half group (F = 14.30, p = 0.0020, ges = 0.12). There were no factors like age, gender, education, handedness, or number of other languages spoken which may have driven this divide between participants.

Comparing Figures 9 and 10 shows that the P3 peak for ideophones in the real condition remains consistent at approximately 5µV. The difference between the two groups is the amplitude of the P3 peak for ideophones in the opposite condition. For the 2AFC top half group, the opposite P3 peak remains consistent, but for the 2AFC bottom half group, the opposite P3 peak rises to 4µV. This means that the greater P3 effect in the test round for participants who scored higher in the 2AFC task is driven by the ERPs in response to ideophones learned in the opposite condition, not ideophones learned in the real condition. Participants in the 2AFC top half group got 88.07% in the real condition and 67.72% in the opposite condition, while participants in the 2AFC low group got 85.15% in the real condition and 75.17% in the opposite condition. This reflects the correlation between sound-symbolic sensitivity and test difference score as shown in Figure 7, and provides a useful marker for more extensive behavioural experiments with a larger sample size.

Discussion

Dutch speakers process Japanese ideophones differently, both behaviourally and neurologically, depending on whether they have learned the words with sound-symbolically matching or mismatching meanings, despite not knowing about the manipulation.

Behaviourally, we found that participants learned the sound-symbolically matching word pairs (i.e. the ideophone and its real translation) better than the sound-symbolically mismatching or non-matching word pairs (i.e. the ideophone and its opposite translation). We also found that, despite doing the learning task, participants were still able to guess the meanings of ideophones at above chance accuracy in a two-alternative forced choice test afterwards. Finally, there was a strong correlation between accuracy and reaction times; the more accurately answered ideophones were answered more quickly. All these behavioural findings closely replicate Lockwood et al. (2016).

In the ERP results, we found no effect of sound symbolism in the learning round, which we speculate is because participants were focused on the learning task; it is possible that effects would arise in a simple judgement or priming task. We did find an effect in the test round, where the presence (or absence) of sound symbolism influenced the amplitude of the P3 and late positivity. The P3 amplitudes per condition did not correlate with participants' test scores per condition, which suggests that the effect is not simply due to ease of learning. However, we did find that the P3 effect magnitude correlated with the behavioural measures of sound-symbolic sensitivity in the 2AFC task performed after the main experiment, which suggests that the P3 effect is related to an individual's sensitivity to sound symbolism.

To further explore this effect, we looked at individual differences between participants and found a relationship between the ERP results and the two behavioural measures of sound symbolism: performance in the sound-symbolic sensitivity check and differences between test scores across conditions. We found that the magnitude of the ERP effects correlated with the performance in the behavioural tasks and thus serves as an index of sound-symbolic sensitivity. This was not hypothesised *a priori,* but the finding provides additional evidence that sound-symbolic sensitivity affects word learning and recognition. It is worth stressing the fact that the behavioural measures from a task measuring sound-symbolic sensitivity predict the ERPs from a completely separate learning and test task (which was done before the participants did the 2AFC task); this suggests that sound-symbolic sensitivity is a consistent process or state which affects how well participants learn sound-symbolic words. To our knowledge, this is the first report of individual differences in sound-symbolic learning and decision tasks being correlated to neurophysiological measures. Rather than noise to be averaged out, these differences can be used to zoom in on the causal processes underlying sound-symbolism and iconicity.

The ERP findings partially mirror existing work on Japanese ideophones and ERPs by Lockwood and Tuomainen (2015), who found that ideophones elicited a greater late positive complex than arbitrary words. Two factors make it difficult to confidently draw functional interpretations for each component from this data; the fact that the P3 and late positive complex are related to all kinds of different functional roles, and the fact that neuroimaging research on sound symbolism in real words is in its infancy. However, we provide two possible interpretations here.

Firstly, the P3 is greater in response to the ideophones learned in the real condition. The P3 is a well-documented component related to attention, and has been functionally separated into a frontal P3a broadly related to stimulus novelty and a parietal P3b broadly related to memory processes (Polich, 2007). The latency and topographic distribution of the effect here suggests that it is a P3b, whose amplitude varies with task demands; increases in memory load reduce P3 amplitude because of the greater task processing demands. Individual difference measures suggest that the P3 effect was related to sound symbolism per condition rather than ease of learning and recognition per condition. Moreover, it appears that the reason for the increased difference in P3b amplitudes between conditions is due to variation in P3b amplitude to ideophones learned in the opposite condition. Participants who scored above the mean in the 2AFC task had a lower P3b amplitude in response to ideophones in the opposite condition; participants who scored below the mean in the 2AFC task had a higher P3b amplitude to ideophones learned in the opposite condition. Coupled with the fact that participants who scored higher in the 2AFC task had a greater accuracy difference between the test scores for ideophones in each condition, this suggests that all participants found learning sound-symbolic words broadly similar, but being more sensitive to sound symbolism makes it harder to learn non-sound-symbolic words and requires extra resource allocation. Therefore, we theorise that the P3 amplitude can be used as an index of the degree to which an individual participant must suppress conflicting cross-modal information during learning and recognition.

In future studies, we would expect to see an even greater P3 amplitude difference between conditions in a similar experiment with pseudowords which deliberately maximise attested cross-modal contrasts. Since eliciting the P3 requires a response during a match/mismatch paradigm (Key, Dove, & Maguire, 2005), it is perhaps unsurprising that we do not find a P3 effect in the initial passive learning rounds.

Secondly, the late positive complex is also greater in response to the ideophones learned in the real condition. The late positive complex in language tasks is generally linked to increased complexity (Lau, Phillips, & Poeppel, 2008), working memory demands (Friederici, Steinhauer, Mecklinger, & Meyer, 1998), or violation of expectation (Van Petten & Luka, 2012), although there is a lot of individual variation (Kos, van den Brink, & Hagoort, 2012). It has also been linked to emotionally arousing stimuli (and referred to as the late positive potential) in non-language ERP literature (Hajcak, Dunning, & Foti, 2009; Hajcak, MacNamara, & Olvet, 2010; Moratti, Saugar, & Strange, 2011). It is possible that the late positive potential observed here and in Lockwood and Tuomainen (2015) are more like those found in the emotion literature. Ideophones are frequently described as being vivid or synaesthetic in how they express meaning, and are particularly well-suited to conveying affective states (Doke, 1948; Kita, 1997). Perhaps the late positive potential elicited by ideophones in Lockwood and Tuomainen (2015) and by ideophones with their real meanings in this study is an indication of their emotional or attentional salience in comparison to arbitrary words or words without sound-symbolic associations. However, there may be a simpler explanation: the strong correlation between P3 effect magnitude and late positive complex effect magnitude ($r=0.46$, $p=0.0124$) suggests that the two components overlap to the extent that the observed late positive complex in this experiment is just a continuation of the large P3 effect, not a separate component reflecting a separate process.

One limitation of the current study is that the stimuli were not counterbalanced across participants. However, we found in pre-tests with the same counterbalanced stimuli in Lockwood et al. (2016) that the behavioural learning effect was consistent for both groups. Another caveat is that the individual difference data is exploratory and should not be taken as conclusive.

Conclusion

Dutch speakers are sensitive to the meanings of Japanese ideophones. Ideophones with their real translations are learned more effectively than ideophones with their opposite meanings due to the congruent cross-modal associations which sound symbolism provides. These associations are accessible despite the learning task, as ideophones were still accurately guessed in a two-alternative forced choice task which took place after the learning task.

Moreover, performance in the 2AFC task actually predicted learning differences between conditions and P3 effect magnitude. This confirms that sound symbolism boosts word learning in adults learning words in a new language, in addition to existing evidence from infants and children as well as adults. It also provides evidence that sound-symbolic cues in Japanese ideophones are available to speakers of an unrelated language, suggesting a fruitful avenue of research into the universality of sound-symbolic cues in ideophones across languages (Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer, in press). While the word learning task is not fully representative of language in a natural context —it is almost impossible to marry full experimental control with full ecological validity— it does go further than the forced choice experiments with pseudowords which make up the majority of sound symbolism research.

Our results pave the way for future work further unravelling the neural correlates and time course of sound symbolism, and suggest that the P3 is heavily implicated in sound symbolism. We suggest that the P3 amplitude is an index of the degree to which the sounds of a word cross-modally match the word's sensory meaning, and that individual differences in sound-symbolic sensitivity constitute a promising inroad for charting the cognitive processes involved in sound symbolism.

References

Akita, K. (2009). A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics. Kobe University.

Diffloth, G. (1972). Notes on expressive meaning. Chicago Linguistic Society, 8, 440–447.

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. Language and Linguistics Compass, 6(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. Trends in Cognitive Sciences, 19(10), 603–615. http://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (in press). What sound symbolism can and cannot do. Language.

Doke, C. M. (1948). The Basis of Bantu Literature. Africa, 18(4), 284–301. http://doi.org/10.2307/3180450

Friederici, A. D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. Biological Psychology, 47(3), 193–221.

Hajcak, G., Dunning, J. P., & Foti, D. (2009). Motivated and controlled attention to emotion: Time-course of the late positive potential. Clinical Neurophysiology, 120(3), 505–510. http://doi.org/10.1016/j.clinph.2008.11.028

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-Related Potentials, Emotion, and Emotion Regulation: An Integrative Review. Developmental Neuropsychology, 35(2), 129–155. http://doi.org/10.1080/87565640903526504

Hamano, S. (1998). The Sound-Symbolic System of Japanese. Center for the Study of Language and Information ; Tokyo.

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How Sound Symbolism Is Processed in the Brain: A Study on Japanese Mimetic Words. PLoS ONE, 9(5), e97905. http://doi.org/10.1371/journal.pone.0097905

Key, A. P. F., Dove, G. O., & Maguire, M. J. (2005). Linking Brainwaves to the Brain: An ERP Primer. Developmental Neuropsychology, 27(2), 183–215. http://doi.org/10.1207/s15326942dn2702_1

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. Linguistics, 35(2), 379–415. http://doi.org/10.1515/ling.1997.35.2.379

Kos, M., van den Brink, D., & Hagoort, P. (2012). Individual Variation in the Late Positive Complex to Semantic Anomalies. Frontiers in Psychology, 3. http://doi.org/10.3389/fpsyg.2012.00318

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. Cognition, 114(1), 19–28. http://doi.org/10.1016/j.cognition.2009.08.016

Kunene, D. (1965). The Ideophone in Southern Sotho. Journal of African Languages, 4, 19–39.

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. Journal of Verbal Learning and Verbal Behavior, 10(4), 427–429. http://doi.org/10.1016/S0022-5371(71)80042-7

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. Nature Reviews Neuroscience, 9(12), 920–933. http://doi.org/10.1038/nrn2532

Levinson, S. C. (2012). The Original Sin of Cognitive Science. Topics in Cognitive Science, 4(3), 396–403. http://doi.org/10.1111/j.1756-8765.2012.01195.x

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. Language Sciences, 1246. http://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-Symbolism Boosts Novel Word Learning. Journal of Experimental Psychology. Learning, Memory, and Cognition. http://doi.org/10.1037/xlm0000235

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. Front. Psychol. Language Sciences, 6, 933. http://doi.org/10.3389/fpsyg.2015.00933

Moratti, S., Saugar, C., & Strange, B. A. (2011). Prefrontal-Occipitoparietal Coupling Underlies Late Latency Human Neuronal Responses to Emotion. The Journal of Neuroscience, 31(47), 17278–17286. http://doi.org/10.1523/JNEUROSCI.2917-11.2011

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. Cognition, 112(1), 181–186. http://doi.org/10.1016/j.cognition.2009.04.001

Oda, H. (2000). An embodied semantic mechanism for mimetic words in Japanese. Indiana University, Indiana.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience, 2011, e156869. http://doi.org/10.1155/2011/156869

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. Frontiers in Psychology, 1(E227). http://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130300. http://doi.org/10.1098/rstb.2013.0300

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology, 118(10), 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. Brain and Language, 128(1), 18–24. http://doi.org/10.1016/j.bandl.2013.11.002

Samarin, W. (1965). Perspective on African ideophones. African Studies, 24(2), 117–21.

Sučević, J., Savić, A. M., Popović, M. B., Styles, S. J., & Ković, V. (2015). Balloons and bavoons versus spikes and shikes: ERPs reveal shared neural processes for shape–sound-meaning congruence in words, and shape–sound congruence in pseudowords. Brain and Language, 145–146, 11–22. http://doi.org/10.1016/j.bandl.2015.03.011

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. International Journal of Psychophysiology, 83(2), 176–190. http://doi.org/10.1016/j.ijpsycho.2011.09.015

Voeltz, F. K. E., & Kilian-Hatz, C. (2001). Ideophones. Amsterdam: John Benjamins Publishing.

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. PLoS Biol, 13(4), e1002128. http://doi.org/10.1371/journal.pbio.1002128

Chapter 6

Synthesised size-sound sound symbolism : initial study and replication.
Lockwood, Hagoort, Dingemanse

Initial study published as:
Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). Synthesized Size-Sound Sound Symbolism. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* [CogSci 2016]. Philadelphia, PA: Cognitive Science Society.

Abstract

Studies of sound symbolism have shown that people can associate sound and meaning in consistent ways when presented with maximally contrastive stimulus pairs of nonwords such as *bouba/kiki* (rounded/sharp) or *mil/mal* (small/big). Recent work has shown the effect extends to antonymic words from natural languages and has proposed a role for shared cross-modal correspondences in biasing form-to-meaning associations. An important open question is how the associations work, and particularly what the role is of sound-symbolic matches versus mismatches. We report on a learning task designed to distinguish between three existing theories by using a spectrum of sound-symbolically matching, mismatching, and neutral (neither matching nor mismatching) stimuli. Synthesized stimuli allow us to control for prosody, and the inclusion of a neutral condition allows a direct test of competing accounts. We find evidence for a sound-symbolic match boost, but not for a mismatch difficulty compared to the neutral condition. We then replicate this finding with twice the sample size.

Introduction

Research into iconicity, where aspects of a word's form reflect aspects of its meaning, has considerably nuanced the classical view of words as wholly arbitrary (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Lockwood & Dingemanse, 2015; Perniss, Thompson, & Vigliocco, 2010). Iconicity is found across languages, both spoken (Dingemanse, 2012) and signed (Emmorey, 2014; Perniss & Vigliocco, 2014), and plays a significant role in language acquisition (Imai & Kita, 2014; Perry, Perlman, & Lupyan, 2015; Yoshida, 2012), language evolution (Cuskley & Kirby, 2013; Verhoef, Kirby, & de Boer, 2015; Zlatev, 2014), and language processing (Lockwood & Tuomainen, 2015; Meteyard, Stoppard, Snudden, Cappa, & Vigliocco, 2015; Westbury, 2005); but it is still unclear exactly how.

Studies have shown that people are sensitive to the meanings of sound-symbolic words in a foreign language, associate certain pseudowords with certain properties depending on their vowels and consonants, and learn new words better when there is a sound-symbolic relationship between form and meaning (Aveyard, 2012; Davis, 1961; Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer, 2016; Kovic, Plunkett, & Westermann, 2010; Lupyan & Casasanto, 2015). The general consensus is that cross-modal correspondences and/or perceptuo-motor analogies between sounds and meanings provide a way of bridging the two domains in consistent ways (Perniss & Vigliocco 2014). However, there is not yet a satisfactory answer to which cross-modal correspondences are implicated in sound symbolism or how exactly these correspondences help people to make mappings. Many experiments have relied on forced choice decisions where participants judge which pseudoword goes with which property (Bremner et al., 2013; Davis, 1961; Köhler, 1929; Nielsen & Rendall, 2011, 2013; Sapir, 1929). This sets up a paradigm where participants consistently identify sound-symbolically matching sets of stimuli (e.g. the pseudoword *bouba* and the round shape, the pseudoword *kiki* and the spiky shape). The combined weight of these experiments is an affirmation of the existence and prevalence of sound symbolism. However, these studies do not address how the associations affect the participants' choices: does a sound-symbolic match provide a mapping boost helping the participant to choose the matching set of stimuli, or does the sound-symbolic mismatch provide a cue to exclude that set of stimuli, or is it a combination of both? Moreover, it is not always clear whether a mismatch is an actual clash, or whether mismatch is simply taken to mean "not matching".

Other experimental designs suggest that it is not as simple as the two-alternative forced choice literature makes out (Monaghan, Mattock, & Walker, 2012; Westbury, 2005). Rating experiments which vary sound-symbolic representations of size along a graded scale have shown that people judge sound symbolism in a graded fashion rather than simply as being there or not (Thompson & Estes, 2011). A graded model of sound symbolism is more detailed, but leaves the same question open: is it driven equally at both ends of the graded spectrum? Learning experiments have shown that it may be one end of a graded spectrum which drives sound-symbolic associations, such as an association between labial sounds and roundness creating an incidental association between non-labial sounds and spikiness (Jones et al., 2014). While it appears that the spiky—round spectrum does not map directly onto the labial/voiced—non-labial/voiceless spectrum suggested by two-alternative forced-choice studies, it remains to be seen whether this imbalance holds for other domains. Finally, other learning experiments suggest there is a sound-symbolic processing bias, but that it is weak and can be overcome with training (Nielsen & Rendall, 2012).

We ran a similar learning experiment with Japanese ideophones (Lockwood, Dingemanse, & Hagoort, 2016) rather than pseudowords. In this study, we taught the ideophones to a group of Dutch participants with no knowledge of Japanese. For half the ideophones, the participants learned the real Dutch translations (e.g. *dik*, or fat, for *bukubuku*, which means fat); for the other half, the participants learned the opposite Dutch translations (e.g. *verdrietig*, or sad, for *ukiuki*, which means happy). In a recognition task, participants remembered the ideophones in the real condition far better than the ideophones in the opposite condition (86.1% recognition accuracy vs. 71.1%). When we repeated the experiment with a set of arbitrary adjectives and another group of participants, there was no sound-symbolic effect across the two conditions (79.1% recognition accuracy in the real adjective condition, 77% in the opposite adjective condition). This is in line with the pseudoword studies that show a mapping boost for sound-symbolically matching stimuli and a mapping difficulty for sound-symbolically mismatching stimuli, although it is not possible to say whether the effect is driven by one or both of these mapping strategies.

In another sound symbolism study with real words, Nygaard et al. (Nygaard, Cook, & Namy, 2009) found a different result. Participants learned Japanese words with their real translations and their opposite translations equally well, but learned words with random translations less well. They proposed that cross-modal correspondences help sound-to-meaning mappings for both matching and mismatching words, as antonym pairs are conceptually very close. Under this interpretation, sound symbolism in learning tasks is not a graded effect. Rather, the lack of any sound-to-meaning correspondence makes word learning harder than having a mismatching or counterintuitive cross-modal clash to build upon.

While using real words from real languages overcomes some of the ecological validity problems of pseudowords, there are other confounds which cannot be completely be ruled out. Firstly, sound-symbolically congruent and incongruent prosody has been shown to affect meaning judgement (Nygaard, Herold, & Namy, 2009). It is possible that our Dutch participants were just picking up on the prosody of the Japanese ideophones rather than the sounds themselves. Secondly, orthography is a constant confound in tasks with both pseudowords and real words (Cuskley, Simner, & Kirby, 2015).

This paper builds on Lockwood et al. (2016) by creating pseudowords in the shape of Japanese ideophones, synthesising the sound stimuli, and limiting the meanings to size. This lets us investigate a spectrum of sound-symbolically matching, mismatching, and neutral stimuli. Here, we take neutral to mean that a relation that is neither an obvious match nor an obvious mismatch. The use of a speech synthesiser to generate the sounds removes prosodic markers from natural speech which may indicate a sound-symbolic relationship. Keeping translations to "big" and "small" lets us work within a well-attested sound-symbolic framework where participants' subjective ratings are highly predictable.

Including a neutral condition while ensuring that the mismatch condition is a cross-modal clash (rather than just a lack of cross-modal correspondence) allows us to adjudicate between different theoretical accounts for sound-symbolic effects. If the participants learn matching pseudowords better than neutral pseudowords, but there is no difference between neutral and mismatching pseudowords, this is evidence for a sound-symbolic match boost as in Lockwood et al. (2016) and Jones et al. (2014). If participants learn matching pseudowords better than neutral pseudowords and neutral pseudowords better than mismatching pseudowords, this is evidence for the graded sound-symbolic rating effect as in Nielsen and Rendall (2012) transferring to sound-symbolic learning. Finally, if participants learn the neutral pseudowords worse than both the matching and mismatching pseudowords, this is evidence for cross-modal correspondences boosting learning regardless of whether the associations correspond or clash, as in Nygaard et al. (2009).

Methods

In the initial experiment (published as a CogSci paper as Lockwood, Hagoort, & Dingemanse, 2016b), 30 participants learned 36 pseudowords in three learning rounds, and were tested immediately afterwards. The replication experiment repeated the experiment with 60 participants. We first describe the stimuli design and selection.

Stimuli design

We created pseudowords in the typical C_1VC_2V-C_1VC_2V pattern found in Japanese ideophones (Akita, 2011). These pseudowords were deliberately created in order to sound big, neutral, or small, based on attested cross-modal correspondences between sound and size. Big-sounding pseudowords featured voiced stops and mid/low back vowels. Small-sounding pseudowords featured voiceless stops and high front vowels. Neutral-sounding pseudowords featured mid-vowels, and had either all voiced, all unvoiced, or a mix of voiced and unvoiced stops. Table 1 shows the distribution of vowels and consonants used in each word type.

Pseudoword type	Consonants	Vowels
big-sounding	[b] [d] [g]	[a] [o]
small-sounding	[p] [t] [k]	[i] [ʏ]
neutral-sounding	[p] [b] [t] [d] [k] [g]	[ɛ] [ə]

Table 1: sound distributions across pseudoword types

We wrote a Matlab script to generate all possible combinations of words according to this pattern, and this resulted in 192 possible pseudowords like *badobado, gɛpɛgɛpɛ,* and *tipitipi*. We then synthesised the pseudowords using the Dutch voice nl2 from the diphone synthesiser MBROLA (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996). All pseudowords were given the same pitch, vowel durations, and prosodic contours.

Stimuli selection

28 native Dutch speakers listened to each synthesised pseudoword and rated how big the word sounded on a Likert scale of 1-7, where 1 represented really small, 4 neutral, and 7 really big. Participants were also told to indicate whether their rating was influenced by a similar-sounding Dutch word in order to detect lexical confounds. We removed 17 pseudowords where at least four participants indicated that it reminded them of something.

In the remaining 175 pseudowords, participants consistently judged the big-sounding words as big (mean=5.57), the neutral-sounding words as neutral (mean=3.90), and the small-sounding words as small (mean=2.68). This was a highly significant effect according to a one-way ANOVA (F=694.3, $p<0.001$), and post-hoc Bonferroni tests showed that this difference was significant between each condition (all ps <0.001). This is shown in Figure 1.

We selected 36 pseudowords for the full experiment according to their mean ratings. For the big-sounding pseudowords, we chose the 12 highest-rated pseudowords; for the small-sounding pseudowords, we chose the 12 lowest-rated pseudowords; and for the neutral-sounding pseudowords, we chose the 12 pseudowords which were rated most closely to 4. All 36 pseudowords were from the originally designated condition, i.e., all 12 big pseudowords were pseudowords which we designed to sound big, and so on.

All pseudowords meant either *groot* (big) or *klein* (small). This set up three conditions: pseudowords that meant big (or small) and sounded big (or small) were sound-symbolically matching, pseudowords that meant big (or small) but sounded small (or big) were sound-symbolically mismatching, and pseudowords that meant big or small but neither obviously matched nor mismatched, thus were neutral. This is illustrated in Table 2.

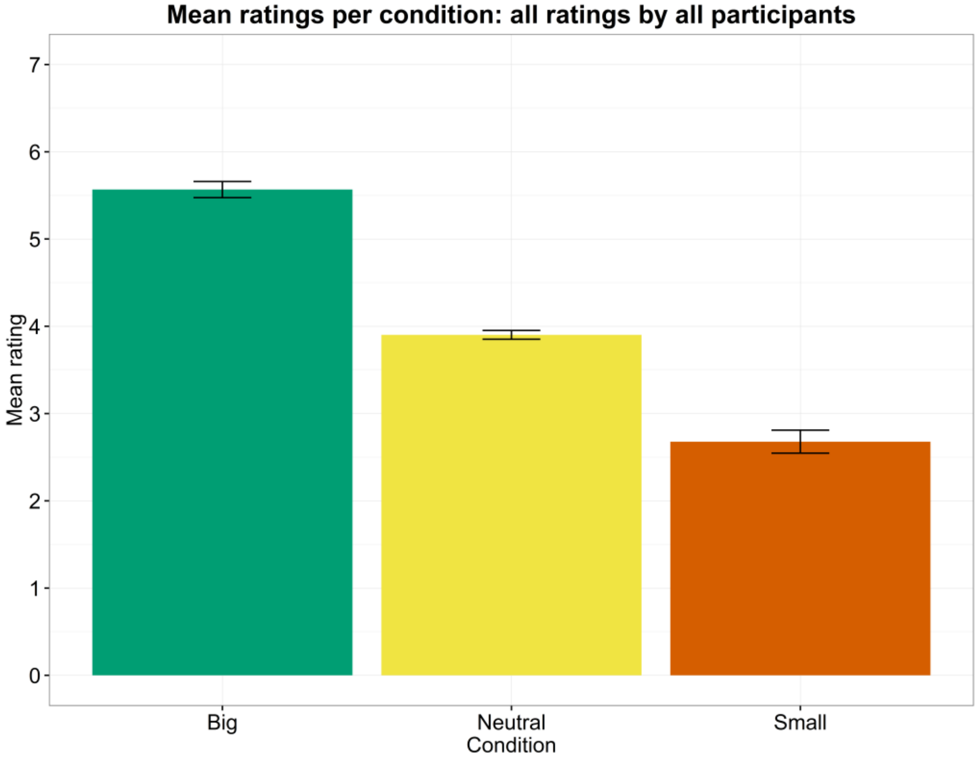


Figure 1: Size ratings per condition

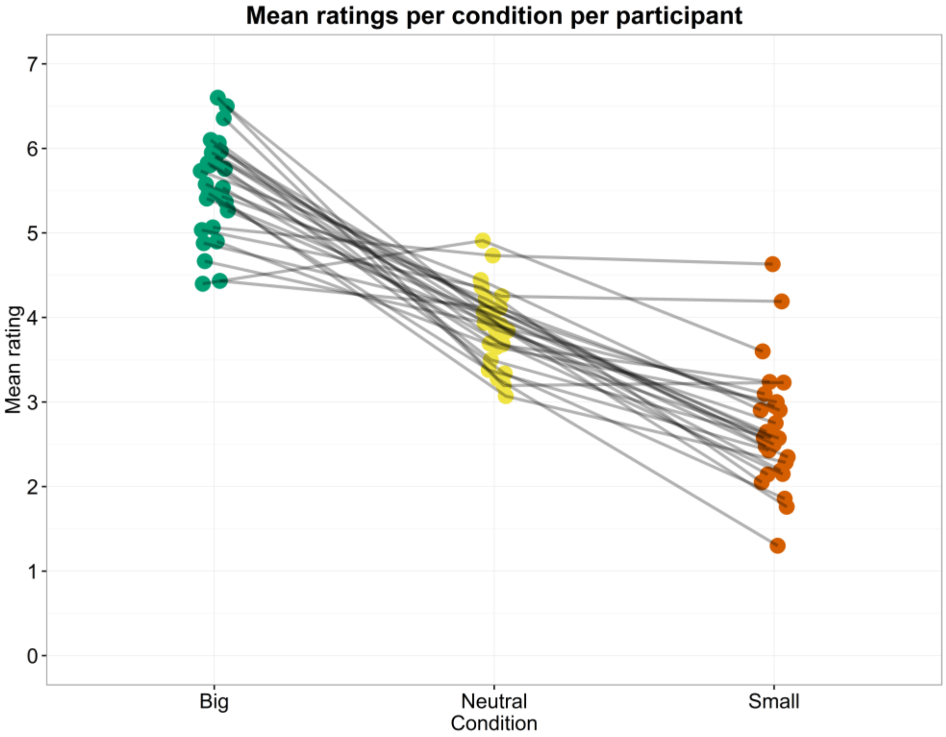


Figure 2: Size ratings per condition per participant. Dots represent individual participants' ratings, and lines between dots show participants in each condition.

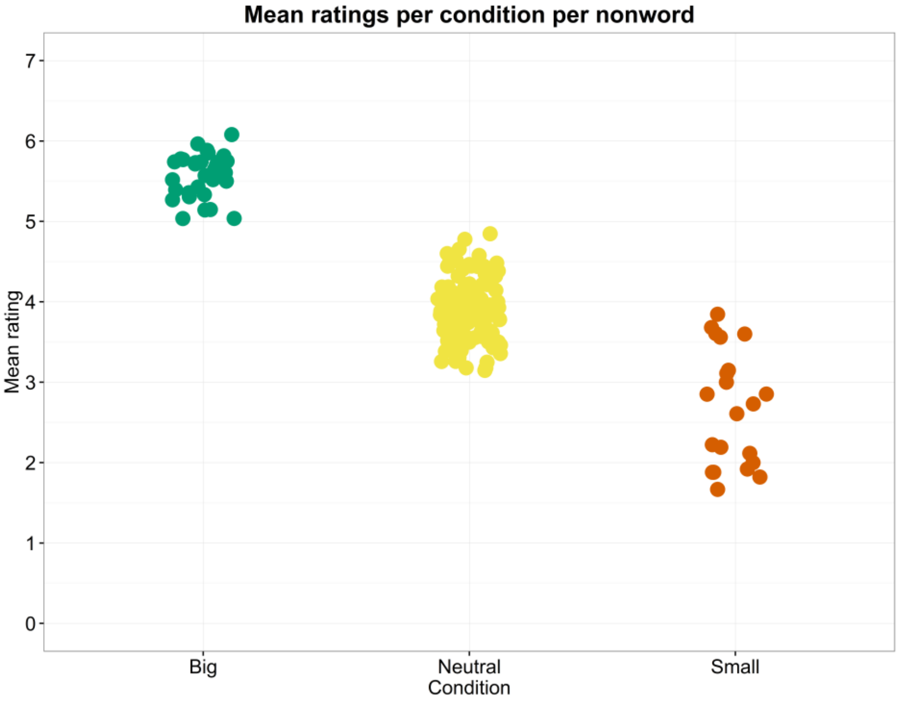


Figure 3: Size ratings per pseudoword

Condition	Pseudoword	Translation
Match	badobado	groot
	tɤpitɤpi	klein
Neutral	detədetə	groot
	gɛpɛgɛpɛ	klein
Mismatch	gogagoga	klein
	tipitipi	groot

Table 2: examples of learning conditions and pseudowords

Initial experiment (n=30)

Participants had three learning rounds in which to learn the pseudowords, and then a test round immediately afterwards. To render the occurrence of multiple forms for a small set of meanings more plausible, they were told that the words came from a language with a complicated adjective agreement system; in a post-test debriefing they were informed of the artificial status of the words. Item translations were counterbalanced across participants. The procedure is illustrated in Figure 4.

We used Presentation to present the stimuli and record responses. In the learning round, the initial Dutch word was presented for 1000ms with 100ms of jitter, followed by a fixation cross for 1000ms with 100ms of jitter. As the pseudoword was played over the speakers, a blank screen was presented for 2000ms with 200ms of jitter. This was again followed by a fixation cross. The final screen with the pseudoword and its Dutch meaning was presented until participants were happy to move onto the next item. Between trials, a blank screen was presented, followed by a fixation cross to announce the beginning of the next trial.

Timings in the test round were identical, except that a question mark was presented instead of a blank screen while the pseudoword played. Participants responded by button press for yes/no answers.

After the test round, participants did a separate two-alternative forced choice task where they guessed the meanings of Japanese ideophones from two antonyms. This was identical to the 2AFC task in Lockwood et al. (2016). The ideophone was played for 2000ms with 200ms of jitter, followed by a fixation cross for 1000ms with 100ms of jitter, followed by a screen showing the Japanese ideophone and the two Dutch options which remained on screen until the participant chose the left or right word by pressing the left or right CTRL key.

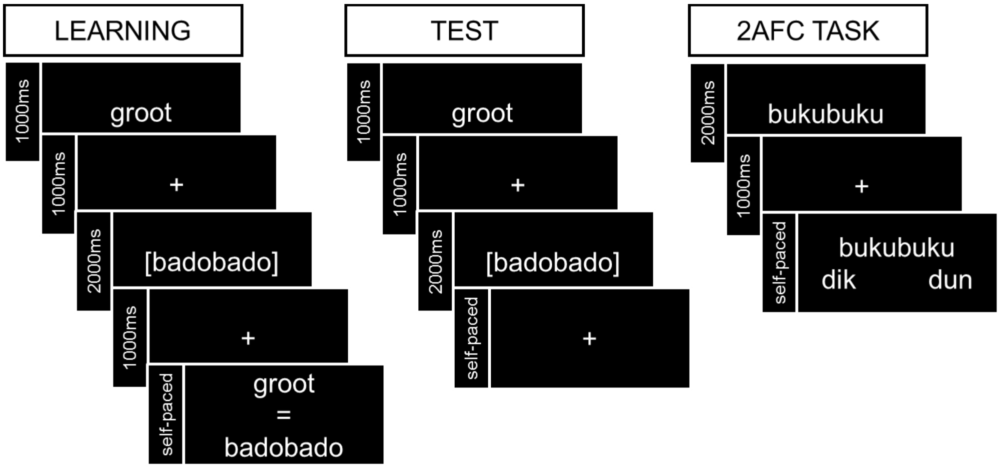


Figure 4: Learning, test round, and 2AFC task procedures.

We tested 33 consenting native Dutch speaking participants (4m, 29f) aged 18-26 (mean: 21y 4m) with normal or corrected-to-normal vision, recruited from the MPI for Psycholinguistics participant database. Three were discarded due to issues with the Presentation script, leaving us with 30 participants in total. This sample size is identical to Lockwood et al. (2016). However, the reduction in the number of items to learn per condition means that more participants are needed to match the power of that study. Therefore, this was intended as an initial experiment to be replicated with a larger sample size.

Replication experiment (n=60)

We ran exactly the same experiment with 62 consenting native Dutch speaking participants (9m, 53f) aged 18-26 (mean: 20y 9m) with normal or corrected-to-normal vision, recruited from the MPI for Psycholinguistics participant database. One was discarded due to not following instructions, and one was discarded due to the computer crashing, leaving us with 60 participants in total. Recruiting male participants was difficult during these experiments. As we had found no gender differences in similar previous studies, we decided that having a sample size large enough to make reliable inferences was more important than having a statistically underpowered but gender-balanced sample size.

In the replication experiment, we also recorded self-paced reaction times for the learning round. Due to a scripting error, we lost the information in the first 14 participants, and so we report learning round reaction time results for participants 15 through 60.

Results

Initial experiment

Participants identified pseudowords in the match condition at 75.56% accuracy, at 66.11% accuracy in the neutral condition, and at 62.50% accuracy in the mismatch condition. This is shown in Figure 5 (plotting all participants separately) and Figure 6 (presenting averages per condition). Error bars in Figure 6 represent standard error.

As the dependent variable was binary—correct or incorrect—we analysed the responses using a mixed-effects logit model with the glmer function of the lme4 (versions 1.1-8) package in R. The data was modelled by including a per-participant and per-pseudoword random adjustment to the fixed intercept with a random slope for the fixed effect by participant. The condition was sum contrast coded to compare match to neutral and neutral to mismatch.

Model comparison between a model with condition as a fixed effect and a model with no fixed effect showed that condition was a significant fixed effect (χ^2=8.36, p=0.015). The best model included a fixed effect of condition, a random effect by participant with random intercepts and random slopes by condition, and random intercept by pseudoword. This model showed that participants did better in the match condition than the neutral condition (β=0.48, SE=0.20, p=0.017), but found no evidence for a difference in performance in the neutral and mismatch conditions (β=-0.11, SE=0.21, p=0.60).

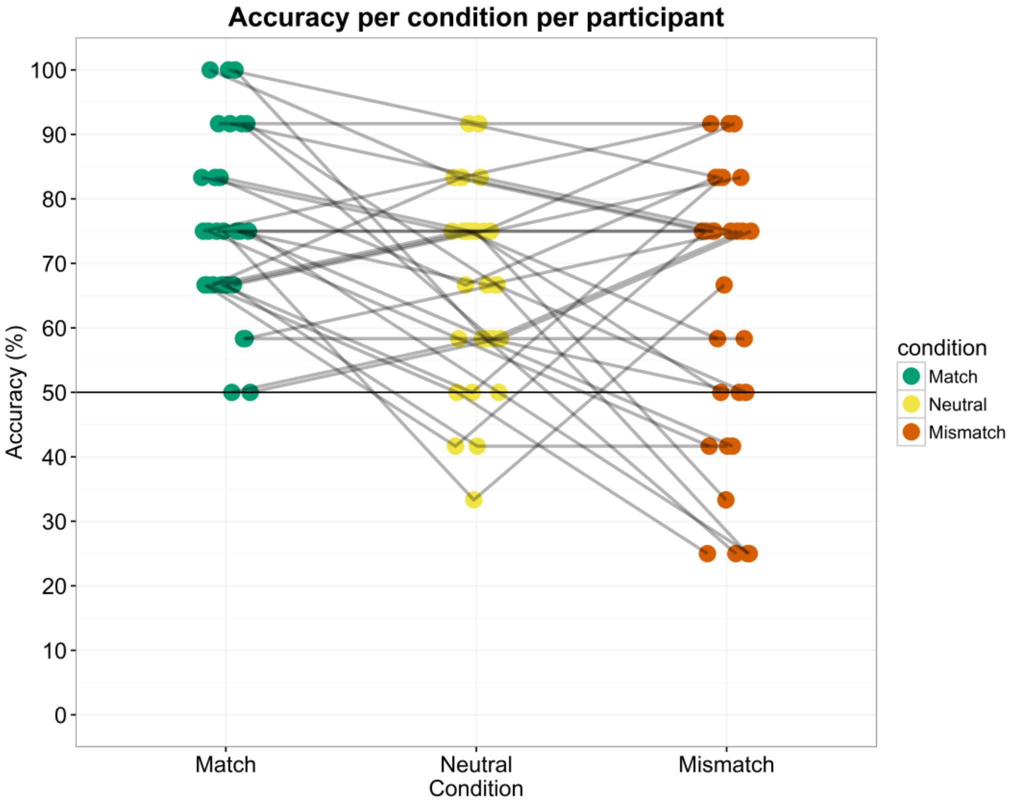


Figure 5: Test round results per participant in the initial experiment (n = 30). Lines connect individual participants' performance across the three conditions.

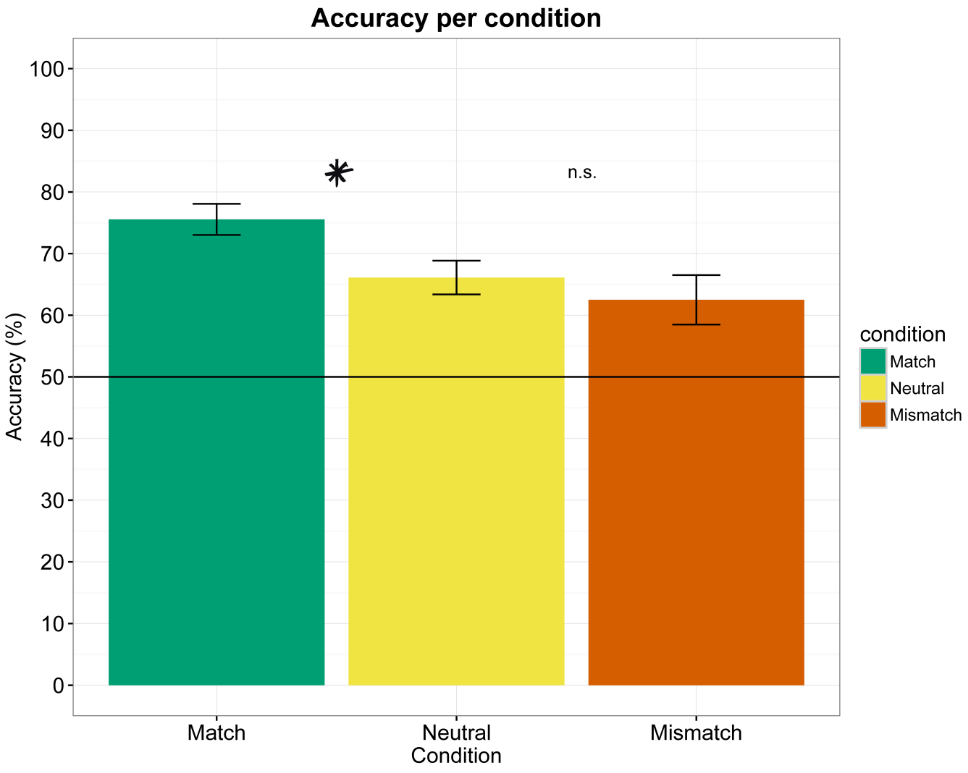


Figure 6: overall test round results

In the 2AFC task, participants guessed the real meaning of the Japanese ideophones as 71.75% accuracy, closely following the 2AFC performance in our previous experiments with these words (Lockwood, Dingemanse, et al., 2016; Lockwood, Hagoort, & Dingemanse, 2016a). We correlated the 2AFC task with the learning task measures to investigate whether 2AFC performance was most obviously linked with participants' sensitivity to a match boost effect or a mismatch difficulty effect.

		r	p
2AFC performance	Overall test score	0.49	0.0058
2AFC performance	Match score	0.15	0.4418
2AFC performance	Neutral score	0.34	0.0655
2AFC performance	Mismatch score	0.49	0.0062

Table 3: correlations between replication experiment test roundand 2AFC task

The 2AFC task scores correlated with the overall test scores and the neutral and mismatch condition scores, but surprisingly not the match condition scores.

Replication experiment

Participants identified pseudowords in the match condition at 77.36% accuracy, at 63.33% accuracy in the neutral condition, and at 56.81% accuracy in the mismatch condition. This is shown in Figures 7 and 8. Error bars in Figure 8 represent standard error.

As above, we analysed the responses using mixed-effects logit modelling with the same per-participant and per-pseudoword random effects, with the data sum contrast coded to compare conditions.

Model comparison between a model with condition as a fixed effect and a model with no fixed effect showed that condition was a significant fixed effect (χ^2=40.24, p<0.001). The best model included a fixed effect of condition, a random effect by participant with random intercepts and random slopes by condition, and random intercept by pseudoword. This model showed that participants did better in the match condition than the neutral condition (β=0.70, SE=0.17, p<0.001), but found no evidence for a difference in performance in the neutral and mismatch conditions (β=-0.28, SE=0.18, p=0.111).

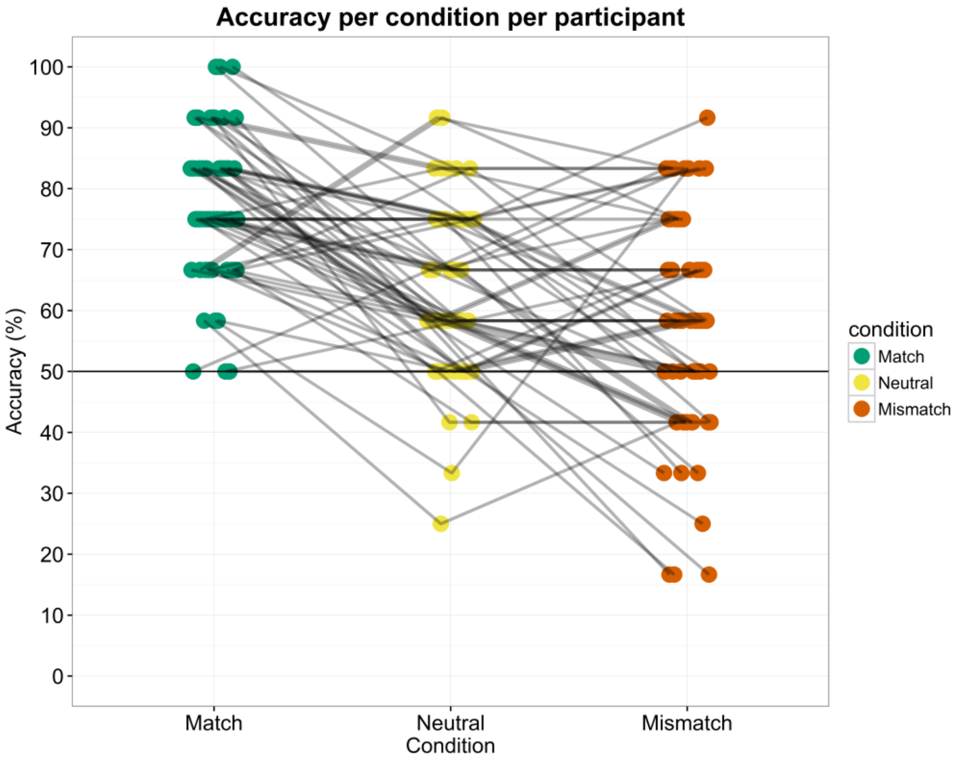


Figure 7: test round results per participant in the replication (n = 60). Lines connect individual participants' performance across the three conditions.

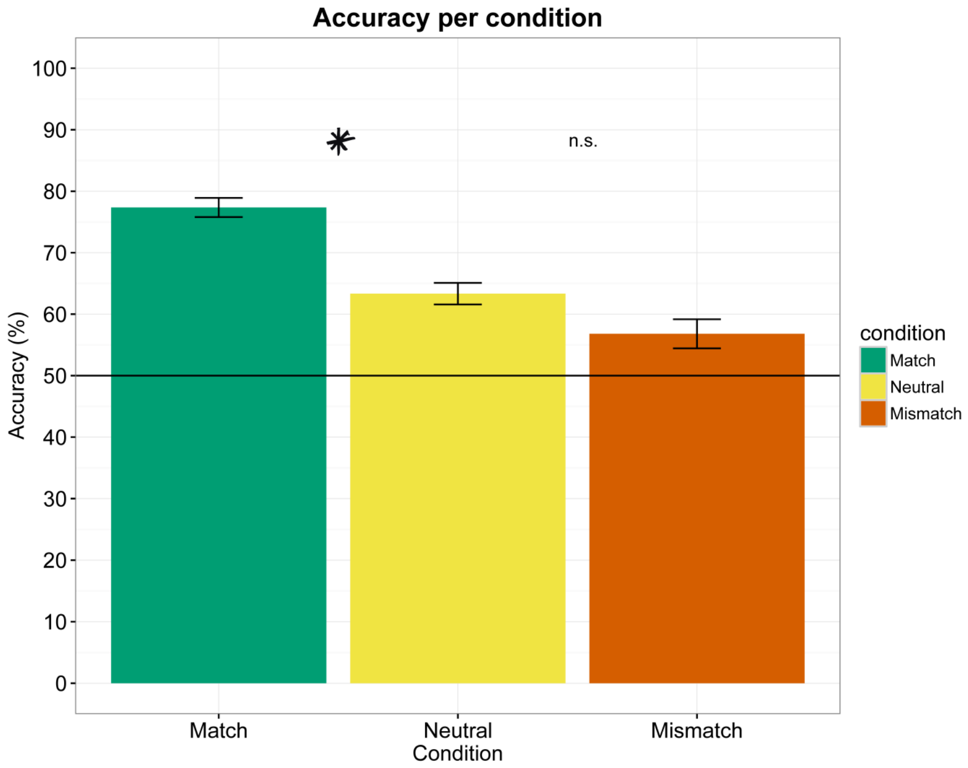


Figure 8: overall test round results

In the 2AFC task, participants guessed the real meaning of the Japanese ideophones at 71.36% accuracy. This time, however, the correlations between the 2AFC task and the learning task were very different.

		r	p
2AFC performance	Overall test score	0.27	0.037
2AFC performance	Match score	0.24	0.0685
2AFC performance	Neutral score	0.02	0.8782
2AFC performance	Mismatch score	0.14	0.2733

Table 4: correlations between replication experiment test round and 2AFC task

The 2AFC task scores again correlated with the overall test scores, but this time, there was no correlation between the 2AFC task scores and the neutral and mismatch condition scores. Moreover, this time there is the hint of a possible correlation between the 2AFC scores and the match condition scores. The lack of consistency in correlations across the initial and replication experiments suggests that we should not read anything into them.

In the replication experiment, we also recorded the amount of time spent on the learning rounds for 44/60 participants (data unavailable for the first 16 participants due to a Presentation scripting error). It is possible that condition differences in the test round scores simply reflect the amount of time spent learning words in those conditions. However, there were no correlations between time spent learning words in a particular condition and test round performance in that condition.

		r	p
Match test score	Match learning time	0.02	0.1903
Neutral test score	Neutral learning time	0.05	0.7281
Mismatch test score	Mismatch learning time	0.08	0.5764

Table 5: (lack of) correlations between replication experiment test round and reaction times

Individual differences

As Figure 5 and Figure 7 show, the relative performance across conditions does not show the same pattern for all participants (this is one reason to show individual data points and not histograms (Weissgerber, Milic, Winham, & Garovic, 2015)). To explore the possibility of individual differences, we classified participants according to a number of performance profiles. The data suggests at least the following distinct performance profiles: (i) graded, where participants' performance in the match condition is better than the neutral condition, which is in turn better than the mismatch condition; (ii) neutral-down, where participants perform worse in the neutral condition than in the match and mismatch conditions; and (iii) other. There were 20 graded participants, 16 neutral down participants, and 24 others. Individual differences in sound symbolism experiments have not yet been widely investigated (Lockwood, Hagoort, et al., 2016a), and the variation in this data shows that more research into individual differences is needed before drawing any definitive conclusions about how sound symbolism works.

Comparisons of the two experiments

Both experiments found evidence for a sound-symbolic match boost effect, while neither experiment found evidence for a difference between the neutral and mismatch conditions.

However, very few of the correlations from the initial experiment were repeated in the full experiment. The one repeated correlation is the one between performance in the 2AFC task and performance in the test in general. This may reflect participants' motivation; participants who did worse in the test were less motivated to take the 2AFC task with Japanese words seriously.

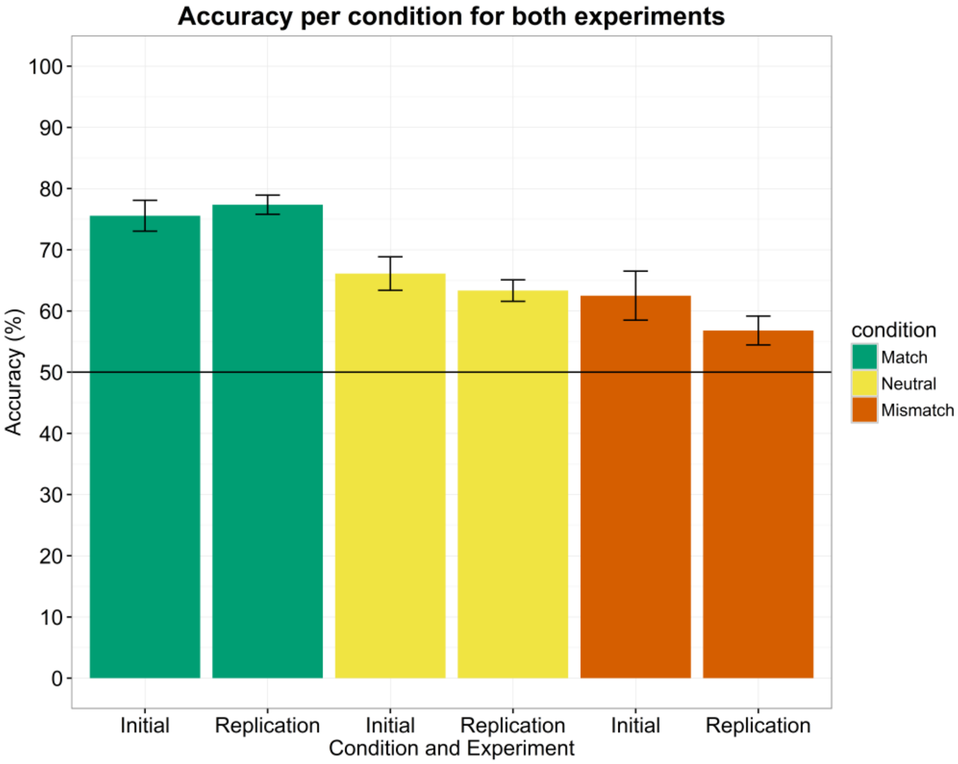


Figure 9: comparison of results from both experiments

Discussion

Sound symbolism research has shown that cross-modal correspondences help people make mappings between sound and meaning. However, it is unclear whether this is because cross-modal correspondences provide a mapping boost or because a lack of a correspondence causes a mapping difficulty. In this study, we build on previous sound-symbolic word learning research by explicitly controlling the type of sound-symbolic relationship in each condition. To zoom in on the question of boost vs. difficulty in a principled, well-controlled way, this study is limited to magnitude symbolism as cued by consonant voicing and vowel quality. It therefore complements other work using real words and a wider range of meanings. Future work may build on these results and extend them into other semantic domains and phonological and prosodic cues.

Participants learned pseudowords which had a variety of sound-symbolic cues to help scaffold word learning. Pseudowords in the match condition had cross-modal correspondences between their sounds and meaning; pseudowords in the mismatch condition had cross-modal clashes between their sounds and meaning; and pseudowords in the neutral condition had neither matching nor mismatching cross-modal information. We then repeated this experiment with twice the number of participants, and found the effects were replicated.

Participants learned the pseudowords in the match condition better than the pseudowords in the neutral condition, but there was no difference in participants' performance in the neutral and mismatch conditions. This suggests that sound-symbolic effects in learning, and perhaps other behavioural tasks, are primarily due to cross-modal correspondences providing a mapping boost. It also suggests that cross-modal mismatches do not provide a mapping boost, but nor do they provide an increased mapping difficulty. This provides initial support for learning experiments suggesting that sound-symbolic bootstrapping depends on the boost effect from matching cross-modal correspondences (Lockwood et al. (2016), Jones et al. (2014), and Imai et al. (2014; Imai, Kita, Nagumo, & Okada, 2008)). It also suggests that the graded perception of sound symbolism in rating tasks (such as in Nielsen & Rendall, 2011; Thompson & Estes, 2011, and indeed, the stimuli selection pre-test for this study) may not extend to a graded learning effect. Finally, it provides some evidence against the idea that any kind of cross-modal associations, whether corresponding or clashing, are better for facilitating sound-symbolic mappings than no cross-modal associations at all. However, this does not rule out the findings of Nygaard et al. (2009). In their experiments, the learning phase was far longer and continued until participants reached a ceiling effect in their accuracy responses. It is possible that there is an initial sound-symbolic match boost during the first stages of word learning, while any kind of cross-modal association can help scaffold word learning during later stages of learning and consolidation.

This study has also tentatively explored individual differences in sound symbolism during learning. This study shows that, overall, participants learned the matching pseudowords better than the neutral pseudowords, while there was no evidence for a difference in how well the participants learned neutral and mismatching pseudowords. This is most obviously shown in Figure 8. However, the dotplots in Figure 7 suggest that it may not quite be so simple. Participants appear to be split, where some (20/60) learn the neutral pseudowords better than the mismatching pseudowords, and some (16/60) learn the mismatching pseudowords better than the neutral pseudowords, while 24/60 participants showed neither a graded nor neutral down effect. On the one hand, this may simply be noise in the overall effect. However, it may also be possible that some participants learn words better when there is a cross-modal association between sound and meaning, whether corresponding or clashing, while other participants may learn in a way that reflects the graded effect of sound-symbolic perception. It is also possible that the participants who showed a graded effect simply hadn't learned the words very well, and were mostly guessing in the learning task according to the sound-symbolic correspondences between the words and the Dutch meanings. Further research on sound symbolism should take possible individual difference measures into account.

In this study, we have addressed the open question in the literature of whether sound symbolism is driven by a cross-modal match boost or a cross-modal mismatch difficulty. Our results show that a cross-modal match between sound and meaning boosts learning across participants, although looking at individual participants suggests that there may be a variety of individual differences and strategies. These individual differences may be more consequential than has previously been assumed, and need more attention in future research. Finally, we have developed a practical, fully-rated, well-balanced set of synthesised stimuli for sound symbolism research.

While this study has highlighted the benefits of using synthesised pseudowords, that doesn't mean that using pseudowords is necessarily better than using real words in sound symbolism research. Studies using pseudowords and real words are complementary, and both are important in the search for the processes and mechanisms underlying sound symbolism.

References

Akita, K. (2011). Toward a Phonosemantic Definition of Iconic Words. In P. Michelucci, O. Fischer, & C. Ljungberg (Eds.), *Semblance and Signification* (pp. 3–18). Amsterdam: John Benjamins Publishing.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & Cognition*, *40*(1), 83–92. http://doi.org/10.3758/s13421-011-0139-3

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). 'Bouba' and 'Kiki' in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165–172. http://doi.org/10.1016/j.cognition.2012.09.007

Cuskley, C., & Kirby, S. (2013). Synaesthesia, cross-modality, and language evolution. In J. Simner & E. M. Hubbard (Eds.), *Oxford Handbook of Synaesthesia* (pp. 869–907). Oxford: Oxford University Press.

Cuskley, C., Simner, J., & Kirby, S. (2015). Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research*, 1–12. http://doi.org/10.1007/s00426-015-0709-2

Davis, R. (1961). The Fitness of Names to Drawings. a Cross-Cultural Study in Tanganyika. *British Journal of Psychology*, *52*(3), 259–268. http://doi.org/10.1111/j.2044-8295.1961.tb00788.x

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, *6*(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, *19*(10), 603–615. http://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What Sound Symbolism Can and Cannot Do: Testing the Iconicity of Ideophones from Five Languages. *Language*, *92*(2), e117-33. http://doi.org/10.1353/lan.2016.0034

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vrecken, O. (1996). The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. In *Fourth International Conference on Spoken Language, 1996. ICSLP 96. Proceedings.* (Vol. 3, pp. 1393–1396 vol.3). http://doi.org/10.1109/ICSLP.1996.607874

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130301. http://doi.org/10.1098/rstb.2013.0301

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130298. http://doi.org/10.1098/rstb.2013.0298

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. http://doi.org/10.1016/j.cognition.2008.07.015

Jones, M., Vinson, D., Clostre, N., Zhu, A. L., Santiago, J., & Vigliocco, G. (2014). The bouba effect: sound-shape iconicity in iterated and implicit learning. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* (pp. 2459–2464). Québec.

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

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19–28. http://doi.org/10.1016/j.cognition.2009.08.016

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Language Sciences*, 1246. http://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-Symbolism Boosts Novel Word Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. http://doi.org/10.1037/xlm0000235

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016a). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016b). Synthesized Size-Sound Sound Symbolism. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society [CogSci 2016]*. Philadelphia, PA: Cognitive Science Society.

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Language Sciences*, 933. http://doi.org/10.3389/fpsyg.2015.00933

Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition*, *7*(2), 167–193. http://doi.org/10.1017/langcog.2014.21

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*. http://doi.org/10.1016/j.neuropsychologia.2015.01.042

Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1152–1164. http://doi.org/10.1037/a0027747

Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*(2), 115–124. http://doi.org/10.1037/a0022268

Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Language and Cognition*, *4*(2), 115–125. http://doi.org/10.1515/langcog-2012-0007

Nielsen, A., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(2), 153–163. http://doi.org/10.1037/a0030553

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*(1), 181–186. http://doi.org/10.1016/j.cognition.2009.04.001

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. *Cognitive Science*, *33*(1), 127–146. http://doi.org/10.1111/j.1551-6709.2008.01007.x

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, *1*(E227). http://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130300. http://doi.org/10.1098/rstb.2013.0300

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition. *PLoS ONE*, *10*(9), e0137147. http://doi.org/10.1371/journal.pone.0137147

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225–239. http://doi.org/10.1037/h0070931

Thompson, P., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *The Quarterly Journal of Experimental Psychology*, *64*(12), 2392–2404. http://doi.org/10.1080/17470218.2011.605898

Verhoef, T., Kirby, S., & de Boer, B. (2015). Iconicity and the Emergence of Combinatorial Structure in Language. *Cognitive Science*. http://doi.org/10.1111/cogs.12326

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PLoS Biol*, *13*(4), e1002128. http://doi.org/10.1371/journal.pbio.1002128

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, *93*(1), 10–19. http://doi.org/10.1016/j.bandl.2004.07.006

Yoshida, H. (2012). A Cross-Linguistic Study of Sound Symbolism in Children's Verb Learning. *Journal of Cognition and Development*, *13*(2), 232–265. http://doi.org/10.1080/15248372.2011.573515

Zlatev, J. (2014). Human Uniqueness, Bodily Mimesis and the Evolution of Language. *Humana.Mente Journal of Philosophical Studies*, *27*, 197–219.

Chapter 7

The all-you-can-rate sound symbolism buffet.
Lockwood, Hagoort, Dingemanse

Abstract

This chapter investigates rating data for three sets of sound-symbolic stimuli: the ideophones from Chapters 4 and 5, the size/sound pseudowords from Chapter 6, and some *bouba/kiki*-esque shape/sound pseudowords. Participants rated the matching stimuli better than the mismatching stimuli in all cases, but the size of the effect outside a two-alternative forced choice task was demonstrably smaller, and there were no ERP effects. This chapter shows that people do clearly make sound-symbolic associations, but that the effects are more subtle than some of the 2AFC literature argues and may only come out in implicit ERP tasks.

Introduction

Previous chapters in this PhD thesis have investigated the main effects of sound symbolism while observing that there may be some interesting individual differences which are obscured by the main effects. In this chapter, participants do three different rating tasks using various different sound-symbolic stimuli. The use of simpler rating tasks also serves as a control experiment for the previous chapters and experiments to explore whether the same stimuli have the same sound-symbolic effects outside the context of a learning task. We used the Japanese ideophones and Dutch translations from the two Japanese ideophone learning chapters in an ideophone rating task, and we used the synthesised pseudowords and Dutch words *groot*/*klein* from the pseudoword learning chapter in a size/sound rating task. Moreover, since no PhD thesis on sound symbolism is complete without a *bouba/kiki* task, we included a shape/sound rating task with adapted stimuli from a previous study by Drijvers et al. (2015).

This study is not simply a final large control experiment. There have been many behavioural sound symbolism studies —see our review paper (Lockwood & Dingemanse, 2015) and the introductions to any other chapter— but as far as we know, there have never been any studies where the same participants do different sound symbolism tasks. As we mention in some of our previous work (Lockwood, Hagoort, & Dingemanse, 2016a, 2016b), interesting individual differences in sound symbolism tasks may often be obscured by the main effects, and far more research into individual differences is needed before creating an overall model of sound symbolism. It is not even known whether individuals are sensitive to all types or merely some types of sound symbolism, nor whether if an individual is, say, particularly good at a *bouba/kiki* task if they are particularly good at guessing the meanings of Japanese ideophones. Up until now, sound symbolism research has shown that people are sensitive to cross-modal correspondences between sound and meaning without being able to compare performances across tasks; having the same participants complete different sound symbolism rating tasks allows for the comparison of different types of sound symbolism.

Behaviourally, our hypotheses were that sound-symbolic effects would arise in each of the three tasks, confirming the large body of sound symbolism literature. We predicted that:

- Japanese ideophones and their real translations would be rated as a better match than Japanese ideophones and their opposite translations
- Synthesised size/sound pseudowords in the matching condition would be rated higher than the neutral condition, which would be rated higher than the mismatching condition
- Fully-matching shape/sound pseudowords would be rated higher than fully-mismatching pseudowords
- Participants would rate congruent sound-symbolic pairs (ideophone and real translation, size/sound pseudoword and match condition, shape/sound

> pseudoword and full match condition) and incongruent sound-symbolic pairs (ideophone and opposite translation, size/sound pseudoword and mismatch condition, shape/sound pseudoword and full mismatch condition) equivalently across all tasks; i.e. that participants' sound-symbolic sensitivity would be similar regardless of the type of sound symbolism

This chapter also addresses the much scantier sound symbolism EEG/ERP literature. Despite the proliferation of behavioural experiments using sound-symbolic stimuli, there are no more than a handful of EEG/ERP studies on sound symbolism (two of which are mine). Most of the published studies out there involve relatively involving tasks, and the only passive task was done with infants (Asano et al., 2015).

Study	Task	Finding
Kovic, Plunkett, & Westermann (2010)	Pseudoword and picture learning	congruent > incongruent (neg peak 140-180ms) at O1 and O2 some N400 effects
Lockwood & Tuomainen (2015)	Sentence reading with unrelated sentence judgment task	increased P2 and late positive complex for ideophones > arbitrary words
Egashira, Choi, Motoi, Nishimura, & Watanuki (2015)	Not actually clear. Also only 4 mimetics vs 4 arbitrary words.	onomatopoeia < common words in early (200-500ms) and middle (500-900ms) LPCs at Pz, P1, P2, POz
Asano et al. (2015)	Passive viewing/listening	increased gamma in match cf mismatch in 1-300ms bigger N400 to mismatch > match 350-550ms, central
Sučević, Savić, Popović, Styles, & Ković (2015)	Lexical decision	40-80ms congruent words positive, congruent pseudowords negative, left fronto temporal 100-160ms incongruent more negative than congruent frontal and Pz 280-320ms congruent more positive than incongruent, Pz 400-620ms words more positive than pseudowords, all over
Lockwood et al. (2016a)	Learning	P3 and LPC difference, all over but especially posterior, real > opposite amplitude diff correlates with individual ppt SS sensitivity
Peeters (2016)	Lexical decision	N2: control words more negative N400: control words more negative LPC: onomatopoeia more positive

Table 1: summary of EEG/ERP sound symbolism studies

As this summary shows, the existing sound symbolism EEG/ERP literature is both small and disjointed. Different ERP effects surface depending on the task, and the only study with a simple sound symbolism paradigm — does this sound go well with this image? — was done with infants.

This may indicate a “file drawer problem”, where research has been done but the results remain in the researcher’s file drawer (or, more accurately, on their harddrive). Given the extent of the behavioural literature, it is surprising to see very few EEG/ERP sound symbolism studies out there. Perhaps sound symbolism is still too small a subfield to have had much experimental interest, and these studies represent the extent of ERP research on the topic. However, my suspicion is that researchers have done EEG/ERP experiments with sound-symbolic stimuli in relatively simple or passive tasks, found no differences between conditions, and that these null results remain unpublished and undiscussed. Accordingly, we hypothesised that there would be no differences between conditions in the ERP data from these experiments.

Methods

Task

There were three tasks, and the order was counterbalanced across participants. In all tasks, participants had to rate how well an auditory item (an ideophone or a synthesised pseudoword) went with its possible referent (a written Dutch word or a round/spiky shape) on a scale of 1 to 7.

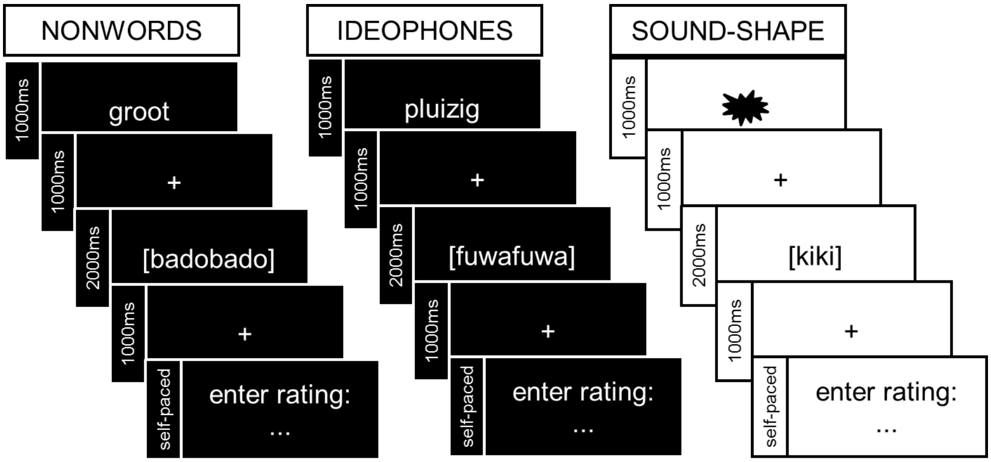


Figure 1: task procedure and timings

Stimuli

The stimuli in the ideophones task were identical to the earlier chapters. There were 38 ideophones with their real and opposite Dutch translations. Participants heard each ideophone twice, once with its real translation, once with its opposite translation. This meant that there were 38 trials in each condition.

The stimuli in the size/sound pseudowords task were identical to the earlier chapter. There were 36 pseudowords; 12 which sounded big, 12 which sounded neutral, and 12 which sounded small. Participants heard each pseudoword four times, twice with the word *groot*, twice with the word *klein*. This meant that there were 48 trials in each condition.

The visual stimuli in the shape/sound task were the same as those used in Drijvers et al. (2015). There were 40 round shapes and 40 pointy shapes, which were created as follows:

> "We wrote a Matlab script that generated random points on a 1000 by 1000 grid, with an adjustable density parameter to generate a predefined number of points. Then, one or multiple of the generated random points were selected, based on an object parameter that could be set to a certain number of figures that needed to be generated and an edging parameter where the number of edges could be predefined. The centre of gravity of the generated figures was based on the mean of the figure. This was done to ensure that all the figures were generated in the middle of the grid. This method ensures the randomized generation of a wide range of properly balanced, contrasting stimuli, constructed according to the best practices outlined in recent work (Nielsen & Rendall, 2011; Westbury, 2005). We made sure all of the figures were single Gestalt shapes, uniformly filled and without holes. All figures had a non-significant difference in the number of pixels to rule out size-effects on sound-symbolic associations and were presented in black to avoid colour confounds." — Drijvers et al. (2015)

In this experiment, we used different pseudowords from Drijvers et al. (2015). Firstly, they were synthesised rather than recorded. Secondly, they were based on the stimuli in Cuskley et al. (2015). They developed a set of consonant contrasts to look at whether shape/sound symbolism for consonants is based on the orthographic shape of the consonants. In their stimuli, all vowels were [e], but we used [i] and [o] so that we can investigate the separate role of vowels and consonants.

This set up a balanced contrast to look at the roles of vowels and consonants separately, and to look at orthographic versus voicing associations. In the literature, voiceless consonants are held to be perceived as spiky while voiced consonants are supposed to be round. However, Cuskley et al. (2015) argue for an orthographic effect, where the shape of the letters dictates the sound-symbolic associations. Under a voicing account, the consonants v and z would be treated as round, while under an orthography account, the consonants v and z would be treated as spiky.

Orthographic				Voicing			
Round V&C	Round V Spiky C	Spiky V Round C	Spiky V&C	Round V&C	Round V Spiky C	Spiky V Round C	Spiky V&C
gogo	koko	gigi	kiki	gogo	koko	gigi	kiki
dodo	toto	didi	titi	dodo	toto	didi	titi
fofo	vovo	fifi	vivi	vovo	fofo	vivi	fifi
soso	zozo	sisi	zizi	zozo	soso	zizi	sisi

Table 2: summary of stimuli and conditions. Vowel and consonant abbreviated to V and C.

Participants

We tested 30 participants in the first round of testing. Only 22 of the first 30 were useable, so we tested an additional 10, making 40 participants tested in total (14m, 26f. 18-28, mean 21y8m). This was reduced to 31 after excluding participants with EEG recording issues.

EEG Recording

EEG was recorded from 61 active Ag/AgCl electrodes, of which 59 were mounted in a cap (actiCap), referenced to the left mastoid. Two separate electrodes were placed at the left and right mastoids. Blinks were monitored through an electrode on the infraorbital ridge below the left eye. The ground electrode was placed on the forehead. Electrode impedance was kept below 10 kΩ. EEG and EOG recordings were amplified through BrainAmp DC amplifiers with a bandpass filter of 0.016–100 Hz, digitised on-line with a sampling frequency of 500 Hz, and stored for off-line analysis.

ERP Analysis

We first used an ocular correction algorithm in BrainVision Analyzer to reduce disruptions from blinks and eye movements. Then, we used semi-automatic artefact rejection to highlight all segments with activity exceeding ±75 µV. 24 participants had less than 20% artefacts in all conditions in all tasks, so we discarded all highlighted trials for them. For another 7 participants who had between 20% and 40% artefacts in some conditions, we inspected each highlighted trial and accepted or rejected them manually. Recordings from nine participants were so messy as to be unusable.

This left 31 participants in total (10m, 21f. 18-28, mean 21y8m). All analyses in this chapter refer only to them. Across these 31 participants, 91.13% of all trials were kept, with at least 30 trials in each condition of each experiment for each participant.

Results

Behavioural results

In the ideophone rating task, participants gave the ideophones and their real translations a mean rating of 4.83, and the ideophones and their opposite translations a mean rating of 3.50.

As there were only two conditions, we used contrast coding for the mixed model. Model comparison showed that a model with a random effect by ideophone with random intercepts was a better fit than a model without (lld = 36.1, $\chi^2 = 72.3$, df = 1, $p < 0.0001$), so we compared a model including random effects by ideophone with the null model to look at the main effect of condition.

Model comparison showed that sound-symbolic condition was a significant main effect (lld = 35.8, $\chi^2 = 71.73$, df = 1, $p < 0.0001$). The model's intercept (i.e. the estimate of the ratings for the opposite condition) was 3.496, and it estimated that ideophones with their real translations were rated 1.34 points higher (se=0.081, t=16.54, p<0.0001).

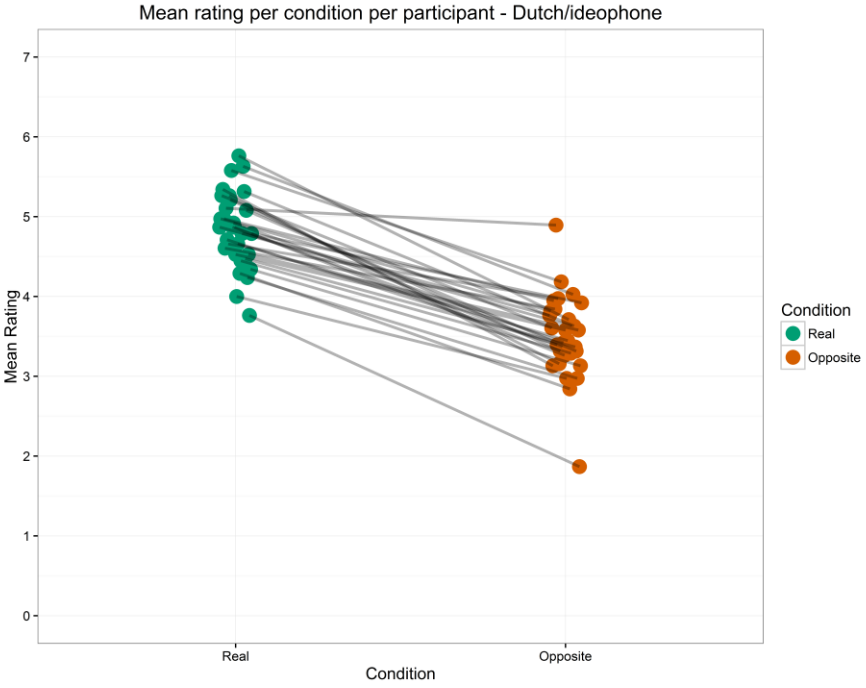


Figure 2: ideophone ratings per condition per participant

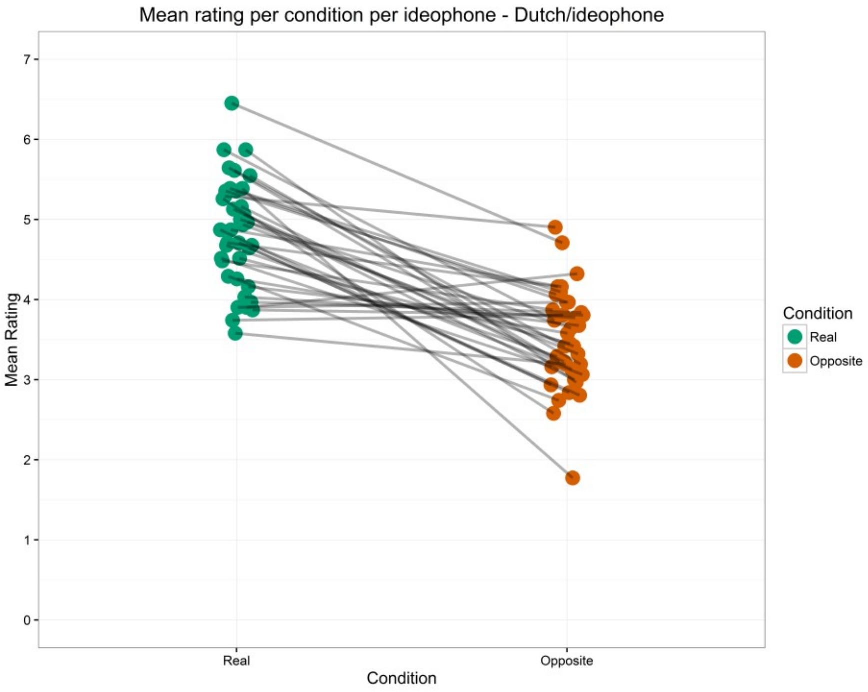


Figure 3: ideophone ratings per condition per ideophone

In the size/sound pseudoword task, participants rated items in the match condition higher than items in the neutral condition, and items in the neutral condition higher than items in the mismatch condition. The mean rating for each condition was 5.63 for the match condition, 3.89 for the neutral condition, and 2.54 for the mismatch condition.

The condition was sum contrast coded to compare match to neutral and neutral to mismatch, since we hypothesised a gradient effect. model comparison showed that a model with a random effect by pseudoword with random intercepts was a better fit than a model without (lld = 12.5, $\chi^2 = 24.95$, df = 1, p < 0.0001), so we compared a model including random effects by pseudoword with the null model to look at the main effect of condition.

Model comparison showed that sound-symbolic condition was a significant main effect (lld = 35.8, $\chi^2 = 71.47$, df = 2, p < 0.0001). The model's intercept (i.e. the neutral condition estimate) was 3.89, and estimated that the match condition was 1.73 points higher (se=0.12, t=14.23, p<0.0001) and that the mismatch condition was 1.35 points lower (se=0.12, t=-11.18, p<0.0001).

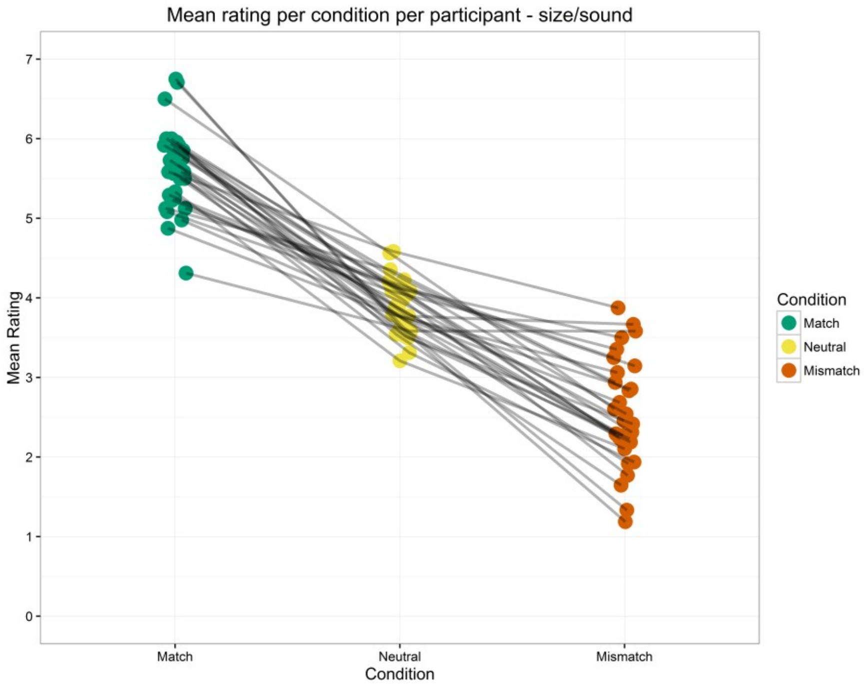


Figure 4: pseudoword ratings per condition per participant

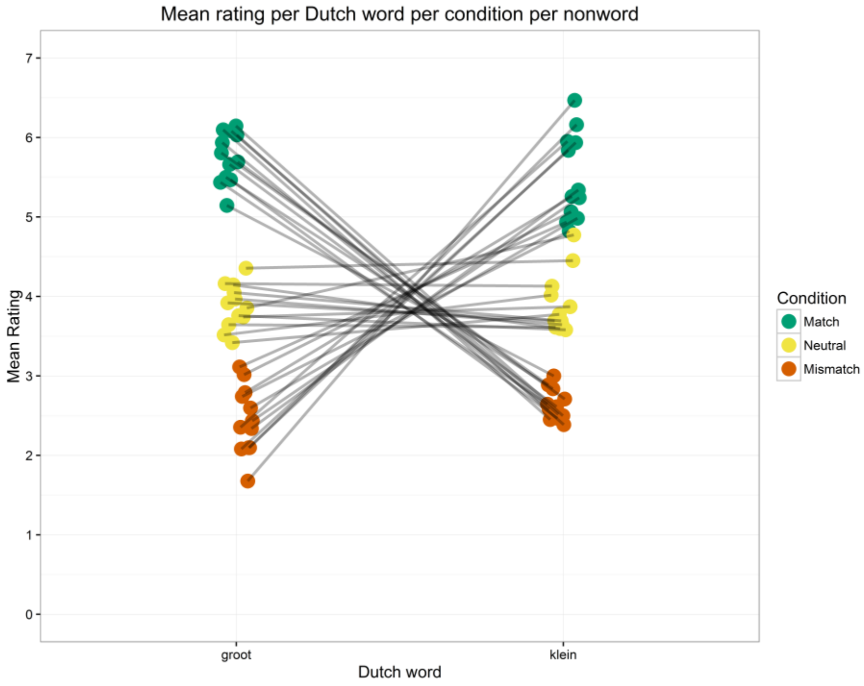


Figure 5: pseudoword ratings per condition per pseudoword

For the shape/sound task, we ran models separately in order to look at orthography and voicing.

In the orthography shape/sound task, participants rated items in the full match condition higher than items in the vowel match condition, items in the vowel match condition higher than items in the consonant match condition, and items in the consonant match condition higher than items in the full mismatch condition. The mean rating for each condition was 4.54 for the full match condition, 4.11 for the vowel match condition, 3.61 for the consonant match condition, and 3.23 for the full mismatch condition.

The condition was sum contrast coded to the vowel match condition, since again we hypothesised a gradient effect. Model comparison showed that a model with a random effect by pseudoword with random intercepts was a better fit than a model without (lld = 14.8, χ^2 = 29.55, df = 1, $p < 0.0001$), so we compared a model including random effects by pseudoword with the null model to look at the main effect of condition. Moreover, model comparison showed that a model adding shape type (i.e. whether the visual stimulus was round or pointy) as a fixed effect was a better fit than a model without (lld = 40.8, χ^2 = 81.70, df = 4, $p < 0.0001$), so we included shape type as a fixed effect.

Model comparison showed that sound-symbolic condition was a significant main effect when compared to a model without sound-symbolic condition (lld = 28.5, χ^2 = 57.11, df = 1, $p < 0.0001$).

The model's intercept (i.e. the vowel match condition estimate) was 3.95, and estimated that the full match condition was 0.43 points higher (se=0.13, t=3.37, p=0.00074), that the consonant match condition was 0.50 points lower (se=0.22, t=-2.25, p=0.025), and that the full mismatch condition was 0.88 points lower (se=0.25, t=-3.59, p=0.00033). The model also estimated that round shapes were rated 0.32 points higher than pointy shapes (se=0.04, t=7.58, $p < 0.0001$), suggesting that round shapes were considered to be more congruent with their associated sounds than pointy shapes.

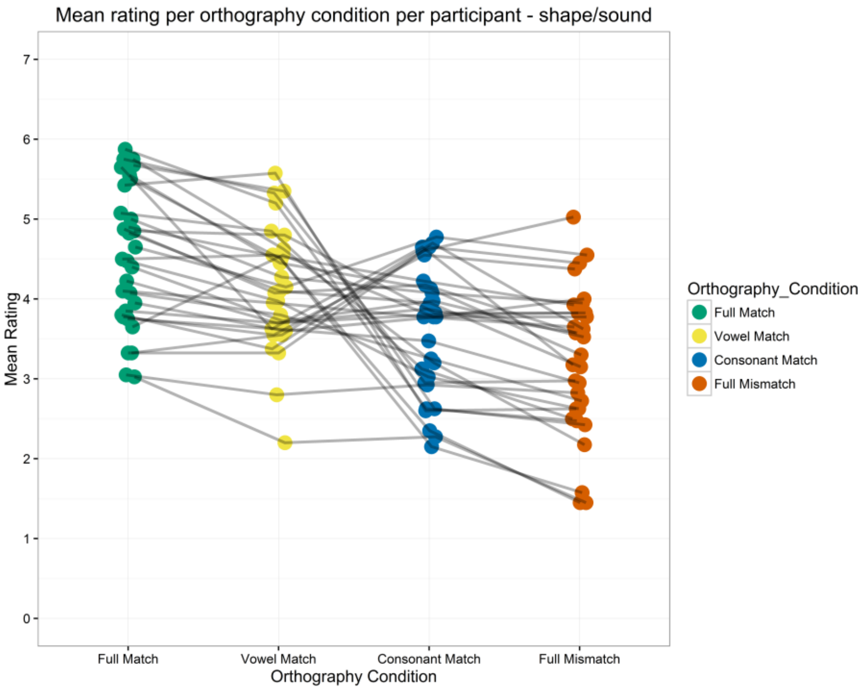


Figure 6: pseudoword ratings per condition per participant in the orthography contrast

In the voicing shape/sound task, participants rated items in the full match condition higher than items in the vowel match condition, items in the vowel match condition higher than items in the consonant match condition, and items in the consonant match condition higher than items in the full mismatch condition. The mean rating for each condition was 4.46 for the full match condition, 4.18 for the vowel match condition, 3.46 for the consonant match condition, and 3.28 for the full mismatch condition.

The condition was sum contrast coded to the vowel match condition, since again we hypothesised a gradient effect. Model comparison showed that a model with a random effect by pseudoword with random intercepts was a better fit than a model without (lld = 14.9, $\chi^2 = 29.76$, df = 1, $p < 0.0001$), so we compared a model including random effects by pseudoword with the null model to look at the main effect of condition. Moreover, model comparison showed that a model adding shape type (i.e. whether the shape was round or pointy) as a fixed effect was a better fit than a model without (lld = 36.5, $\chi^2 = 73.08$, df = 4, $p < 0.0001$), so we included shape type as a fixed effect.

Model comparison showed that sound-symbolic condition was a significant main effect when compared to a model without sound-symbolic condition (lld = 28.5, $\chi^2 = 57.12$, df = 1, $p < 0.0001$).

The model's intercept (i.e. the vowel match condition estimate) was 4.02, and estimated that the full match condition was 0.32 points higher (se=0.13, t=2.10, p=0.035), that the consonant match condition was 0.63 points lower (se=0.22, t=-2.78, p=0.0054), and that the full mismatch condition was 0.90 points lower (se=0.26, t=-3.52, p=0.00044). The model also estimated that round shapes were rated 0.32 points higher than pointy shapes (se=0.04, t=7.58, p < 0.0001), suggesting that round shapes were considered to be more congruent with their associated sounds than pointy shapes.

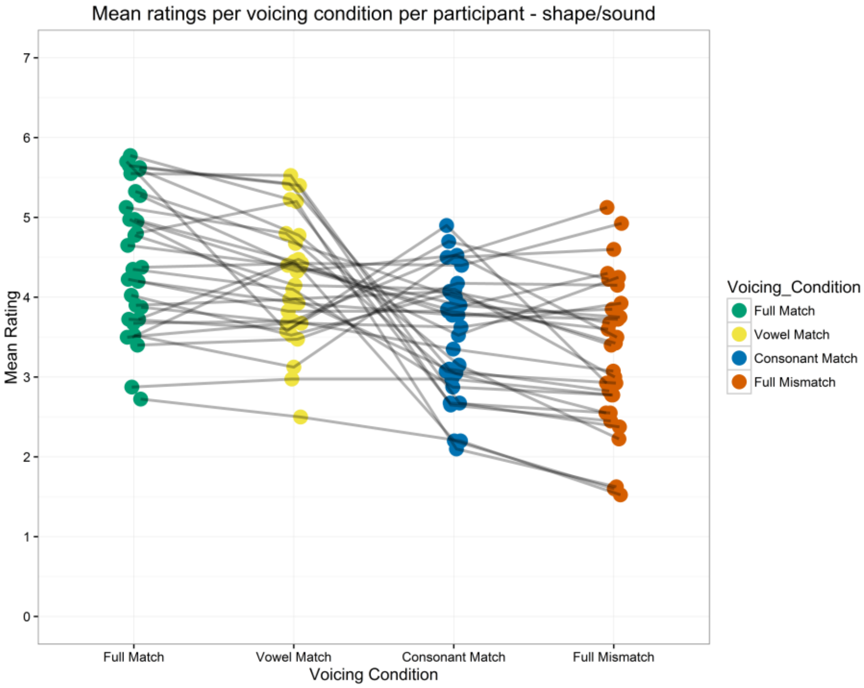


Figure 7: pseudoword ratings per condition per participant in the voicing contrast

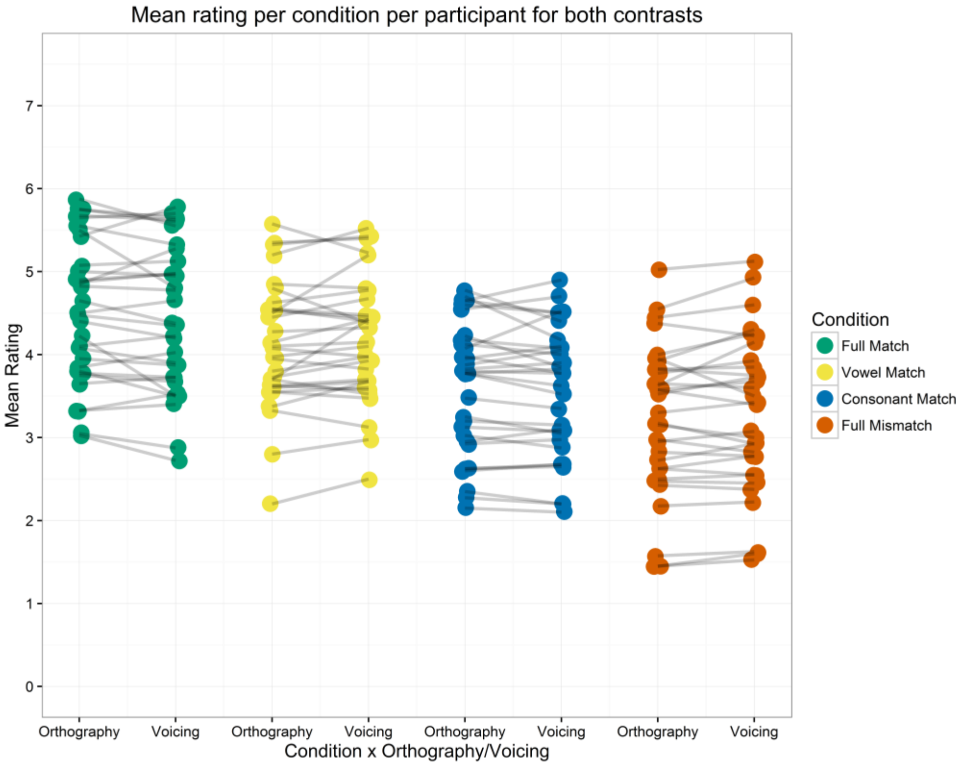


Figure 8: pseudoword ratings per condition per participant in the both contrasts

In both the orthography and voicing contrasts, shape type was found to make a difference, with participants giving higher ratings in any of the match conditions to round shapes than spiky shapes:

Orthography/Voicing	Condition	Round Rating	Pointy Rating
Orthography	Full Match	4.69	4.38
Orthography	Vowel Match	4.44	3.78
Orthography	Consonant Match	3.78	3.43
Orthography	Full Mismatch	3.21	3.24
Voicing	Full Match	4.74	4.18
Voicing	Vowel Match	4.39	3.98
Voicing	Consonant Match	3.75	3.36
Voicing	Full Mismatch	3.24	3.31

Table 3: summary of orthography and voicing contrast results

This is in line with Jones et al. (2014), who found that the cross-modal correspondences between round shapes and round sounds are more salient than those between spiky shapes and spiky sounds.

Correlations across behavioural measures

In order to explore participants' sensitivity across sound symbolism tasks, we correlated the ratings from each task. Correlations were Benjamini and Hochberg-corrected for multiple comparisons. Significant correlations are reported in the table below:

In short, the ideophone results correlate with the pseudoword size/sound results, and the pseudoword size/sound results correlate with the pseudoword shape/sound results, but there's no correlation between the ideophone results and the pseudoword shape/sound results. We also correlated the rating difference between the extreme conditions per participant; i.e. real minus opposite for the ideophone task, match minus mismatch for the size/sound pseudoword task, and full match minus full mismatch for the shape/sound pseudoword task.

first measure	second measure	corrected p value	r value
Ideophone Opposite Rating	Ideophone Real Rating	0.0030	0.58
Ideophone Opposite Rating	Size/Sound Mismatch	0.0028	0.59
Size/Sound Match Rating	Size/Sound Mismatch Rating	0.0020	-0.60
Size/Sound Match Rating	Orthography Full Match Rating	0.012	0.52
Size/Sound Match Rating	Orthography Full Mismatch Rating	0.015	-0.51
Size/Sound Match Rating	Voicing Full Match Rating	0.015	0.51
Size/Sound Match Rating	Voicing Full Mismatch Rating	0.0081	-0.54
Size/Sound Mismatch Rating	Size/Sound Neutral Rating	0.030	0.47
Orthography Consonant Match Rating	Orthography Full Mismatch Rating	< 0.0001	0.88
Orthography Consonant Match Rating	Voicing Consonant Match Rating	< 0.0001	0.97
Orthography Consonant Match Rating	Voicing Full Mismatch Rating	< 0.0001	0.88
Orthography Full Match Rating	Orthography Vowel Match Rating	< 0.0001	0.81
Orthography Full Match Rating	Voicing Full Match Rating	< 0.0001	0.96
Orthography Full Match Rating	Voicing Vowel Match Rating	< 0.0001	0.84
Orthography Full Mismatch Rating	Voicing Consonant Match Rating	< 0.0001	0.87
Orthography Full Mismatch Rating	Voicing Full Mismatch Rating	< 0.0001	0.98
Orthography Vowel Match Rating	Voicing Full Match Rating	< 0.0001	0.85

Orthography Vowel Match Rating	Voicing Vowel Match Rating	< 0.0001	0.94
Voicing Consonant Match Rating	Voicing Full Mismatch Rating	< 0.0001	0.81
Voicing Full Match Rating	Voicing Vowel Match Rating	< 0.0001	0.78
Ideophone Real-Opposite Rating Difference	Size/Sound Match-Mismatch Rating Difference	0.012	0.50
Size/Sound Match-Mismatch Rating Difference	Orthography Full Match-Full Mismatch Rating Difference	0.015	0.47
Size/Sound Match-Mismatch Rating Difference	Voicing Full Match-Full Mismatch Rating Difference	0.021	0.44
Orthography Full Match-Full Mismatch Rating Difference	Voicing Full Match-Full Mismatch Rating Difference	< 0.0001	0.98

Table 4: summary of behavioural correlations

Two correlations from within the same task are of particular interest; the per-participant correlations between real and opposite condition ideophone ratings, and match and mismatch condition size/sound pseudoword ratings. In the ideophone task, there is a positive correlation; participants who gave higher ratings to ideophones in the real condition also gave higher ratings to ideophones in the opposite condition. However, in the size/sound pseudoword task, there is a negative correlation; participants who gave higher ratings to pseudowords in the match condition gave lower ratings to pseudowords in the mismatch condition. Possible reasons for this disparity are outlined in the discussion section.

ERP results

The main waveforms and topographic plots of each experiment are plotted on the following pages. Statistical analyses follow, but it is clear from the waveforms that there is no ERP effect between conditions in any case.

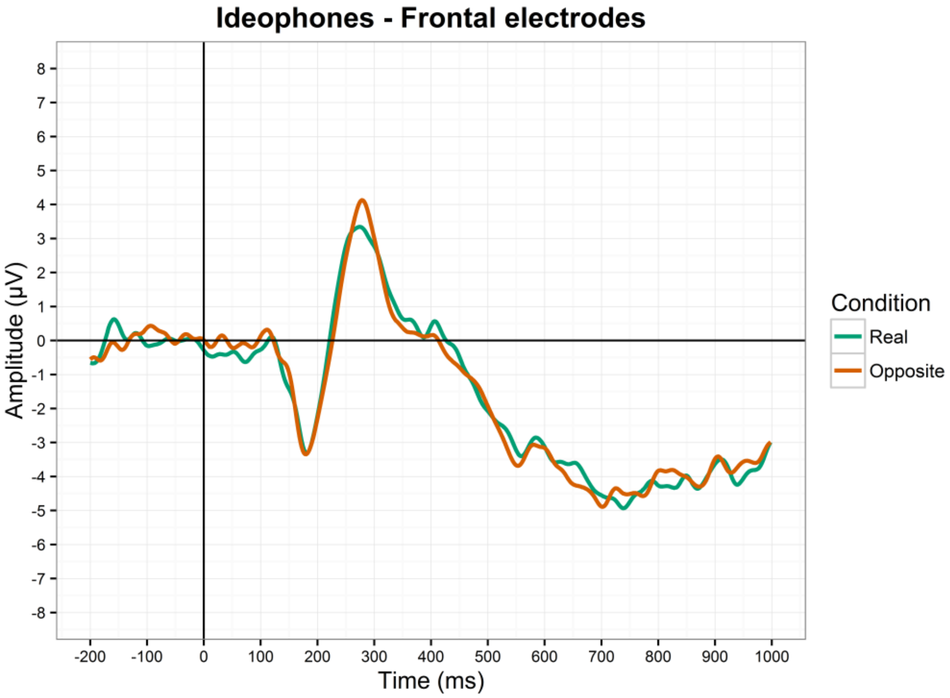


Figure 9: ERPs for the ideophone task at the frontal electrodes

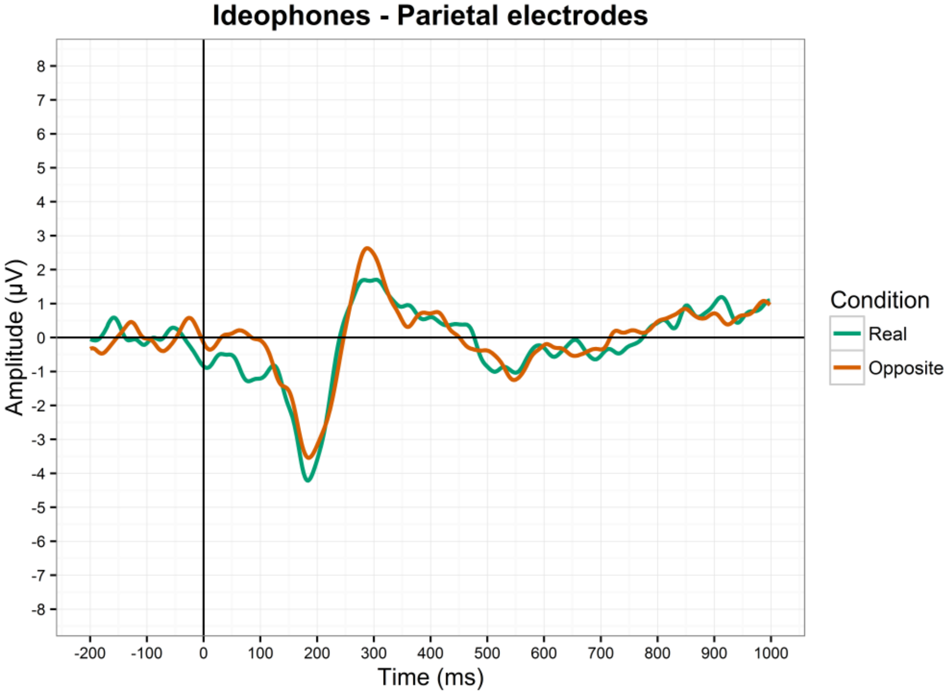


Figure 10: ERPs for the ideophone task at the parietal electrodes

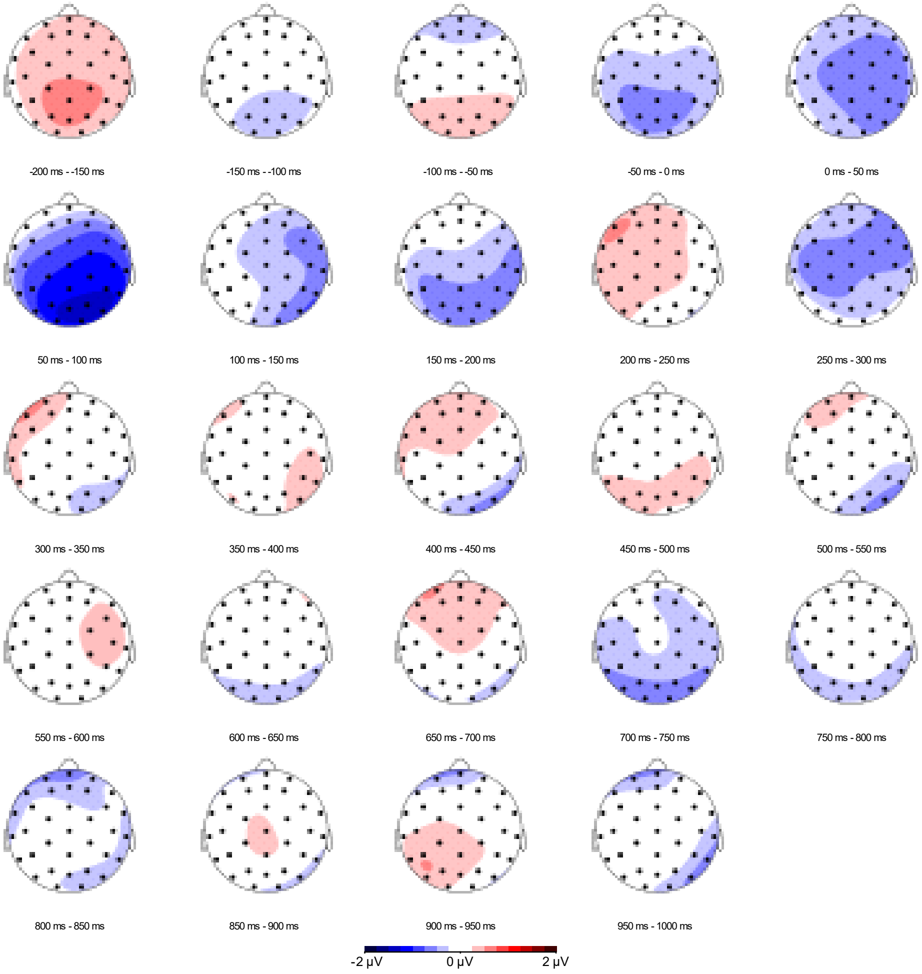


Figure 11: topoplots for the ideophone task (real minus opposite conditions)

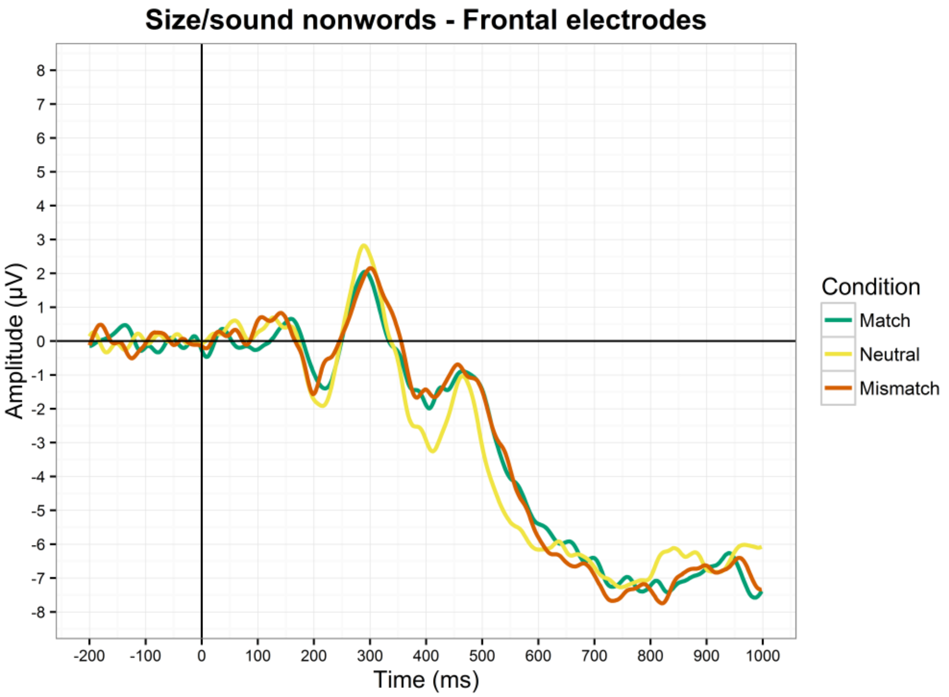


Figure 12: ERPs for the size/sound task at the frontal electrodes

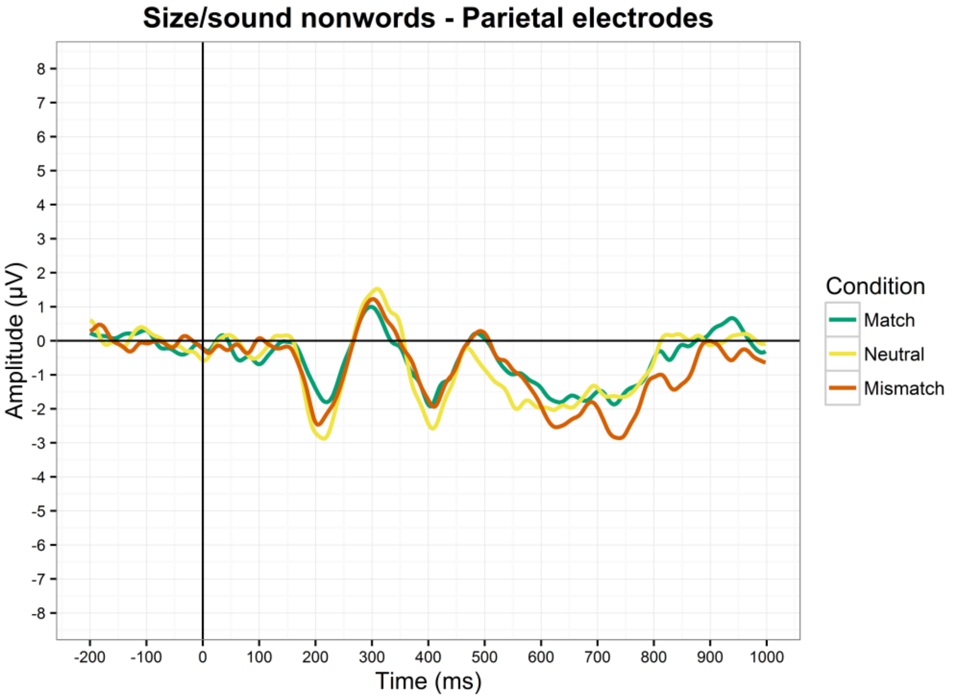


Figure 13: ERPs for the size/sound task at the parietal electrodes

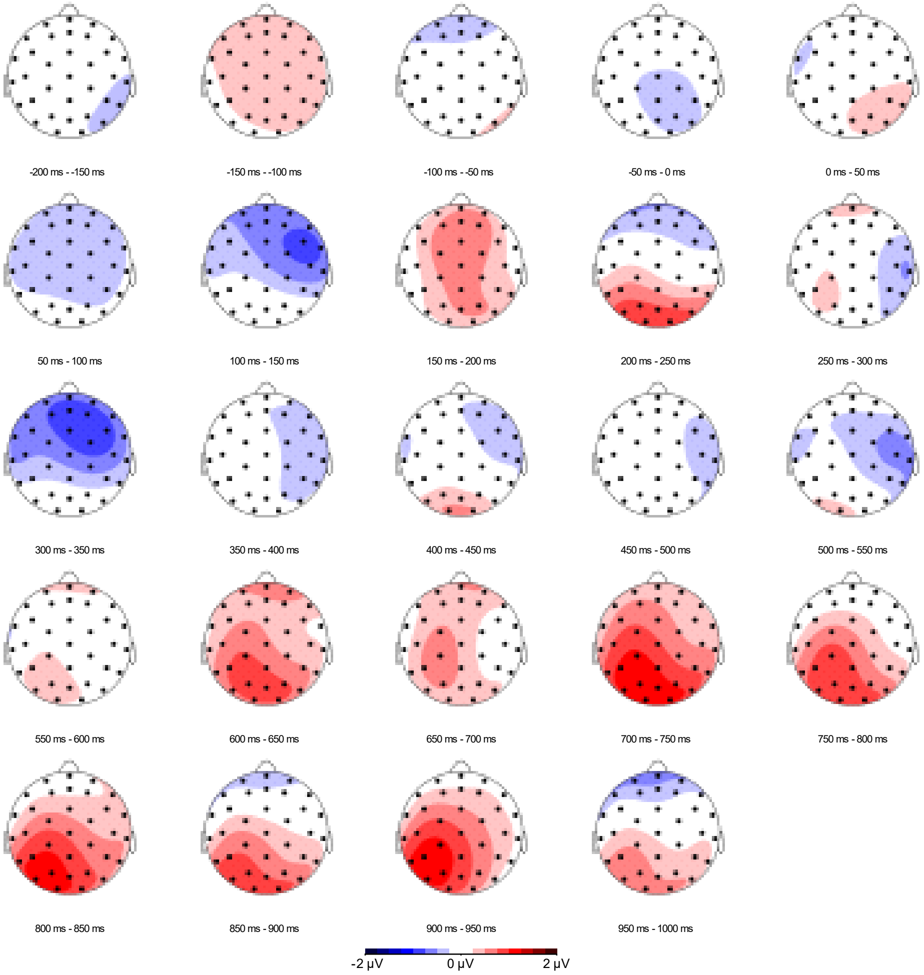


Figure 14: topoplots for the size/sound task (match minus mismatch conditions)

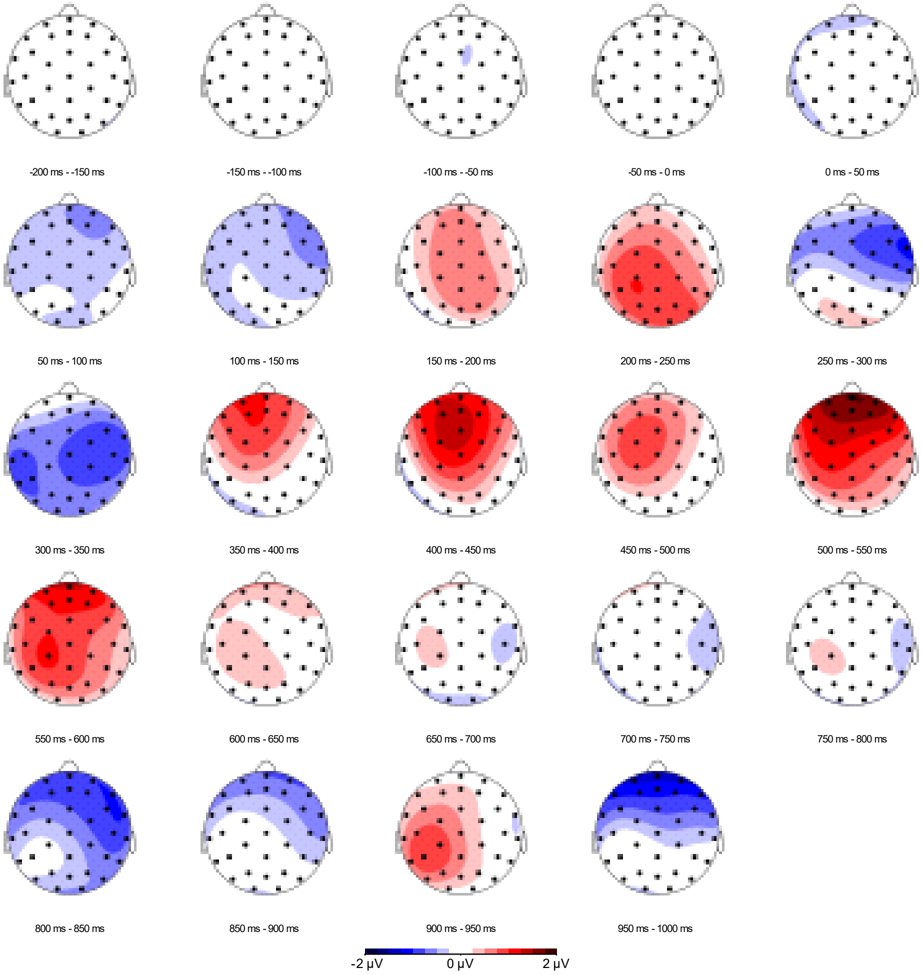


Figure 15: topoplots for the size/sound task (match minus neutral conditions)

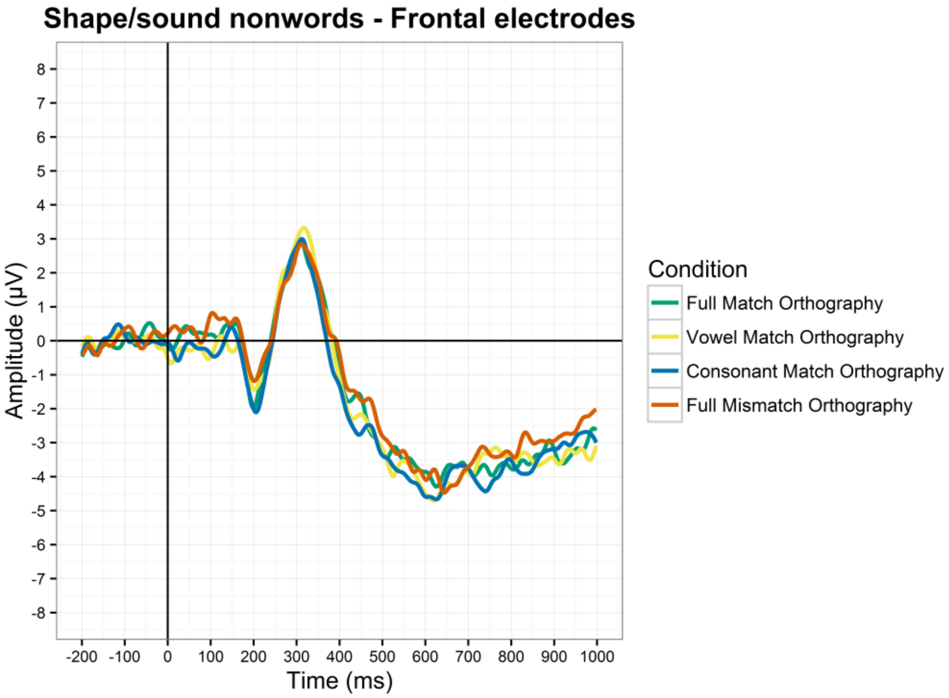


Figure 16: ERPs for the shape-sound task at the frontal electrodes (orthography contrast)

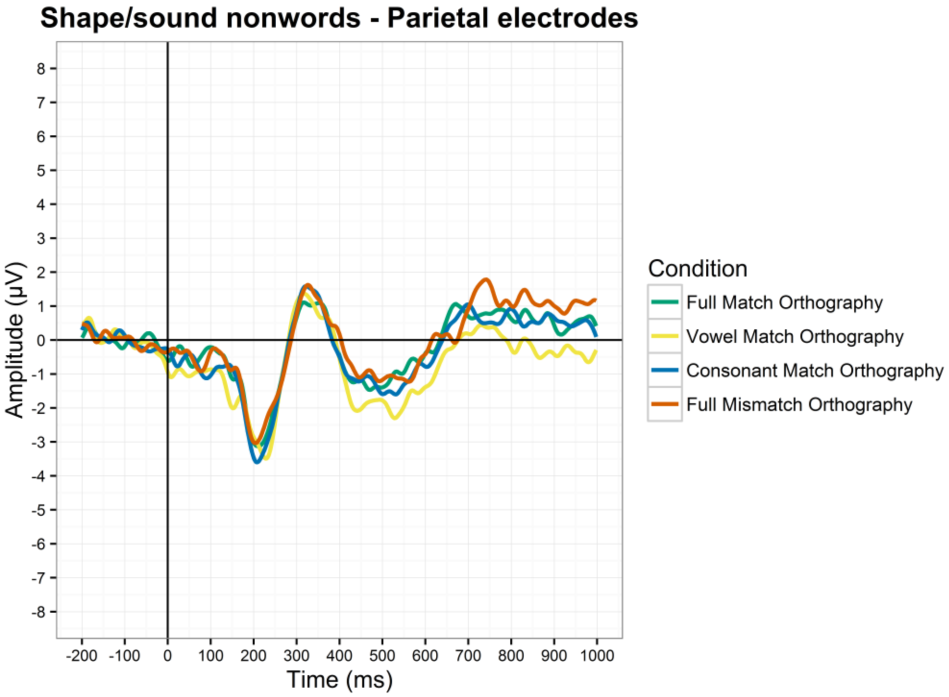


Figure 17: ERPs for the shape-sound task at the frontal electrodes (orthography contrast)

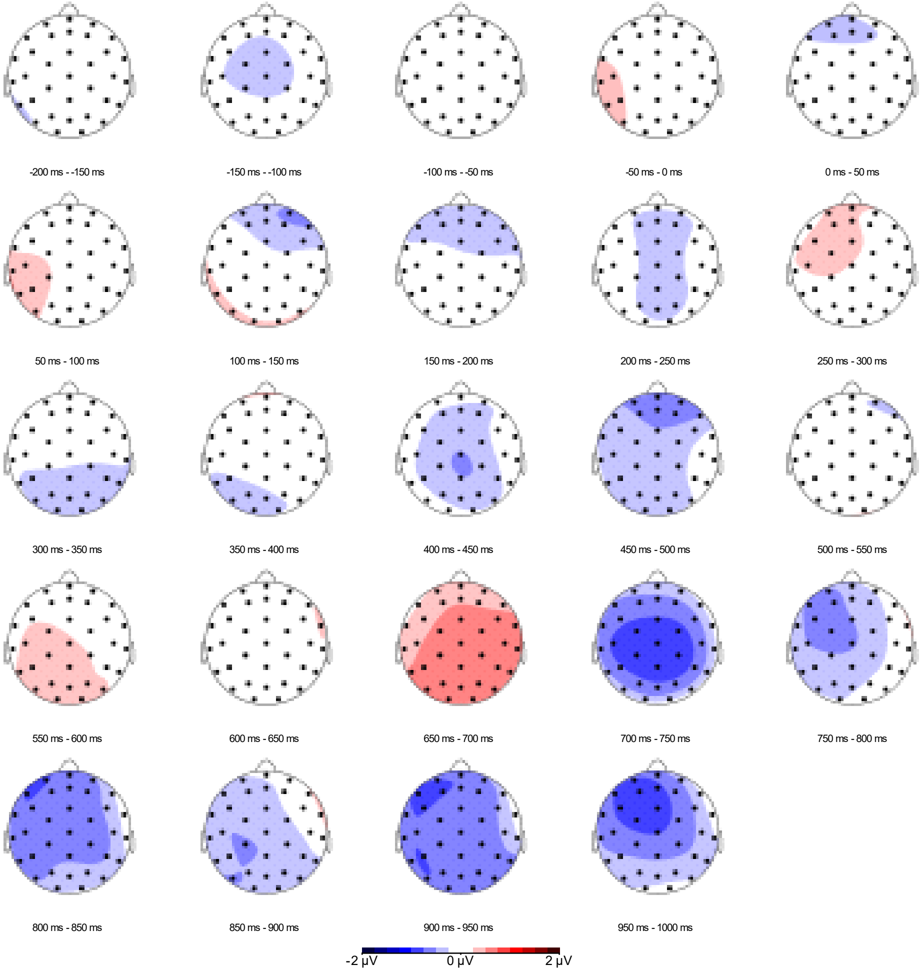


Figure 18: topoplots for the shape-sound task (full match minus full mismatch conditions - orthography)

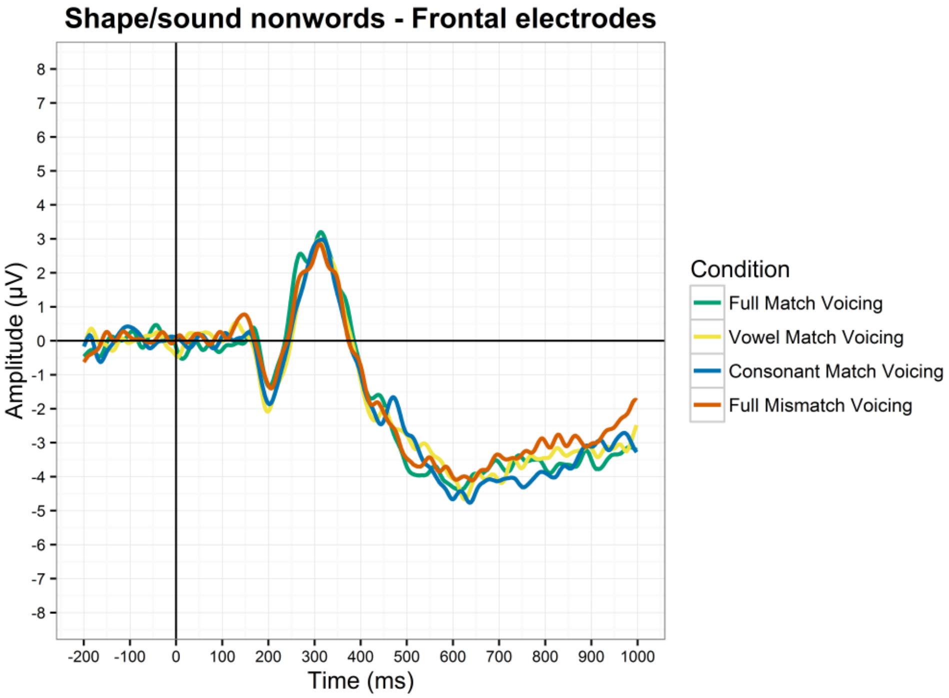


Figure 19: ERPs for the shape-sound task at the frontal electrodes (voicing contrast)

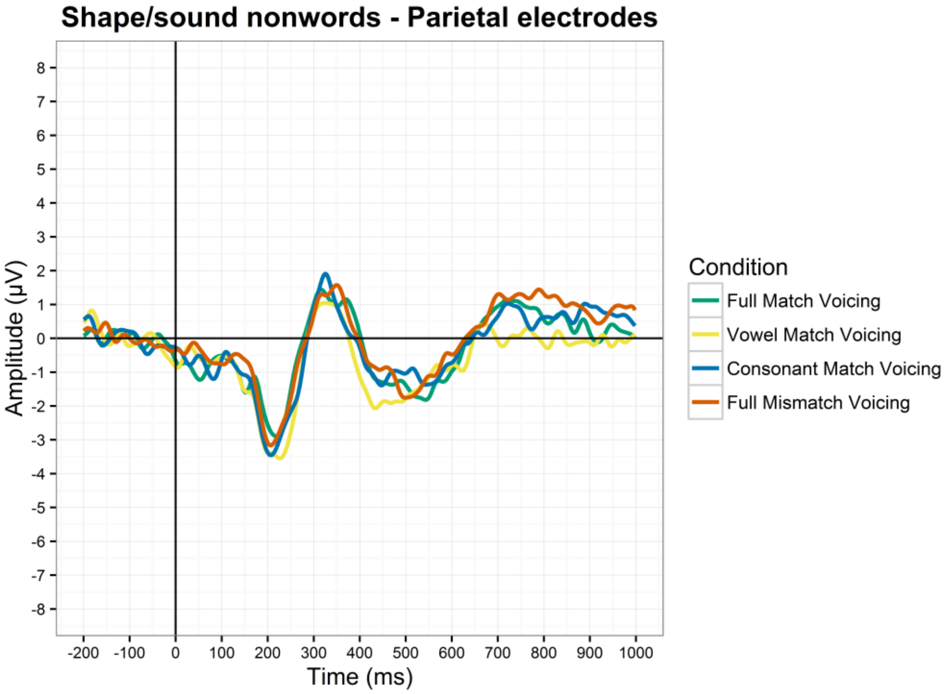


Figure 20: ERPs for the shape-sound task at the frontal electrodes (voicing contrast)

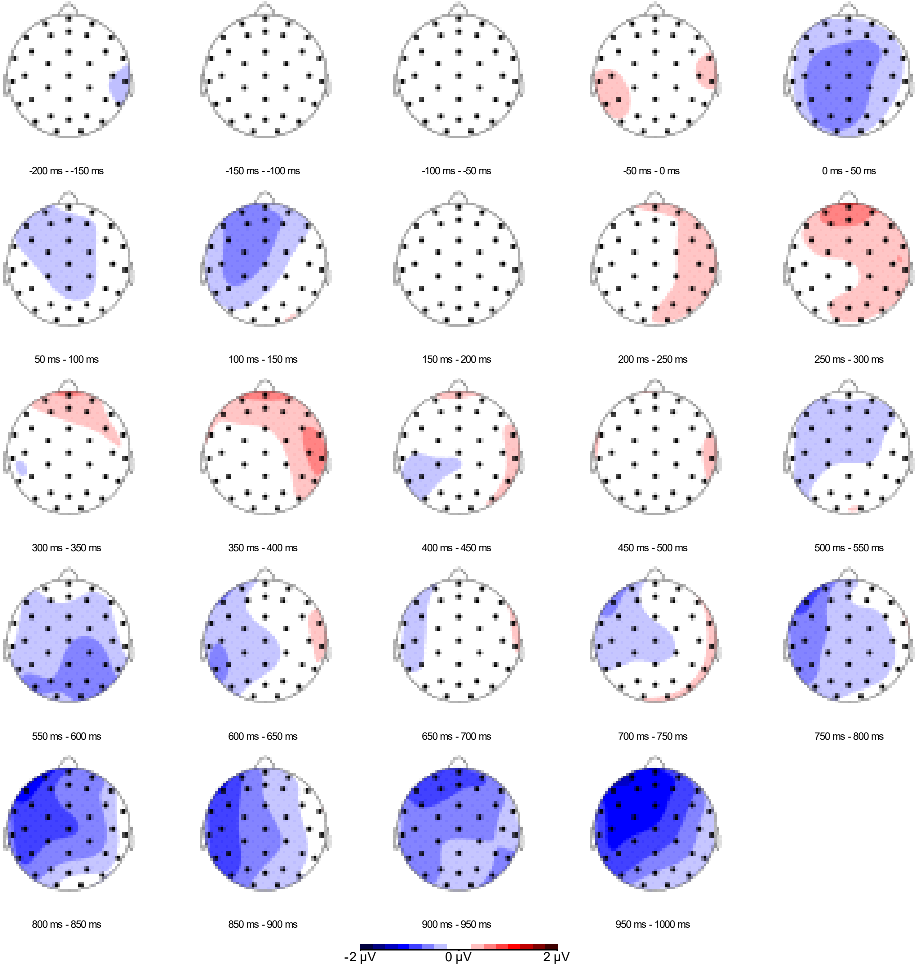


Figure 21: topoplots for the shape-sound task (full match minus full mismatch conditions - voicing)

Visual inspection of the waveforms is enough to show that there is no real difference between any of the conditions. To investigate the results further, we tried the exploratory —and statistically dubious— procedure outlined by Sučević et al. (2015) in their paper, which is also similar to the approach of Kovic et al. (2010). This approach involves performing within-subject condition x quadrant ANOVAs for the mean ERP amplitude in every 20ms time window throughout the epoch, and then considering a significant 20ms time window to be actually significant if the 20ms time windows either side were also significant. For example, if significant effects were found in five time windows from 80ms to 180ms, the window of middle three from 100ms to 160ms would be considered significant. This approach is not ideal, but we use it here to show how there are no effects in the ERP data even when the approach lends itself to false positives.

In the ideophone task, there were 3 significant 20ms windows, detailed in the table below. No significant window was between other significant windows, so there would be no effect even under this method.

Window	F value	P value	"Significant"?
20-40ms	4.21	0.0489	no
60-80ms	7.66	0.00958	no
80-100ms	5.67	0.0238	no

Table 5: ideophone task rolling 20ms ANOVAs

In the size/sound pseudoword task, there were 6 significant 20ms windows, detailed in the table below. Only one significant window (520-540ms) was between other significant windows.

Window	F value	P value	"Significant"?
500-520ms	4.60	0.0138	no
520-540ms	4.28	0.0183	yes
540-560ms	3.80	0.0278	no
740-760ms	3.80	0.0278	no
820-840ms	7.49	0.00125	no
840-860ms	4.49	0.0152	no

Table 6: size/sound task rolling 20ms ANOVAs

In the shape/sound pseudoword task, we ran the procedure separately for orthography and voicing interpretations of sound symbolism. Since there were four conditions, we ran the procedure with all conditions and also with only the two extreme conditions, full match and full mismatch.

In the orthography interpretation, the only individually significant time window was 980-1000ms (F=3.24, p=0.0258), which would not be considered significant under this method. When looking at only full match and full mismatch conditions, there were no 20ms windows which were individually significant at all.

In the voicing interpretation, the only individually significant time window was 260-280ms (F=3.68, p=0.015), which would not be considered significant under this method. When looking at only full match and full mismatch conditions, the only individually significant time window was 940-960ms (F=4.24, p=0.0483), which would not be considered significant under this method.

From this, it is safe to say that performing ANOVAs over larger windows with stricter window selection criteria would prove completely inconclusive.

Individual differences exploration across behavioural and ERP measures

In our earlier research (Lockwood et al., 2016a), we found that splitting the participant group in half according to their behavioural performance showed that the behavioural task predicted the ERP amplitude difference. Participants who were more sensitive to sound symbolism as measured by a 2AFC task to guess the meaning of Japanese ideophones had a much bigger ERP amplitude difference during a separate ideophone recall task. Compared to the group effect, the top half of participants according to 2AFC task performance had a bigger ERP effect; the bottom half of participants according to 2AFC task performance had a much smaller (and in fact non-significant) ERP effect. In this analysis, we divided the participants into two groups by the difference between their ratings in the extreme conditions; i.e. mean real minus mean opposite ratings for the ideophone task, mean match minus mean mismatch ratings for the size/sound pseudoword task, and mean full match minus mean full mismatch ratings for the shape/sound task. Splitting the group in half according to rating difference showed nothing. The waveforms for each group were almost identical to the whole group waveforms (and accordingly, the waveforms are not plotted here so as to save space).

Discussion

In this chapter, we used the same stimuli as in the previous chapters (the JEP:LMC paper, the Collabra paper, and the CogSci paper) and an extra shape/sound task derived from Drijvers et al. (2015) (same shape stimuli) and Cuskley et al. (2015) (similar linguistic stimuli). Instead of learning the pairs and then being tested on them, as in my previous chapters, participants just had to read/see the words/shapes, listen to the sound-symbolic (pseudo)words, and rate how well they went together.

Behaviourally, everything was exactly as expected. Participants gave higher ratings to sound-symbolically congruent pairs than sound-symbolically neutral and/or mismatching pairs. Moreover, this was done in a graded way. This is especially interesting with regards to the size/sound stimuli, which are rated in a graded way and learned in a match boost way (Lockwood et al., 2016b). We also found evidence to support Jones et al. (2014), who found that the cross-modal correspondences between round shapes and "round" sounds are more salient than the correspondences between spiky shapes and "spiky" sounds. This is additional evidence that the *bouba/kiki* effect as found in two-alternative forced choice tests is actually more of a *bouba* effect; the strong roundness correspondences drive the effect and create a partially artefactual tendency to link spiky shapes to "spiky" sounds.

There were also informative correlations across the experiments. Firstly, participants tended to rate the stimuli in the same way, i.e. people who rated matching conditions highly in the size/sound pseudoword experiment also rated the real condition highly in the ideophone experiment, and people who rated matching conditions highly in the size/sound pseudoword experiment also rated the matching condition highly in the shape/sound pseudoword experiment. However, there was no correlation between the ideophone and the shape/sound experiment ratings.

There are two ways in which the shape/sound task is different from the ideophone task. Firstly, the *nature* of the referential stimuli (i.e. the things that the sound-symbolic (pseudo)words iconically depict). The Dutch translations of Japanese ideophones are completely linguistic stimuli, which participants have to read in order to get an idea of their sensory meaning. The spiky or round figures are non-linguistic; while seeing the round or spiky shapes probably activates the linguistic concepts of spikiness and roundness, the stimuli themselves are shapes whose sensory properties are obvious, independent of language (as indeed Asano et al. (2015) found with 11 month old infants). Secondly, the *variety* of the referential stimuli. In the ideophone task, there were as many Dutch translations as there were ideophones, all from a variety of semantic domains, whereas in the shape/sound stimuli, the referential stimuli were limited to a simple dichotomy of round vs. spiky. Given these two major differences in the stimuli in the two tasks, it is perhaps unsurprising that the participants don't perform similarly in the ideophone and shape/sound tasks; both are about sound symbolism, but they are different types of sound symbolism in different contexts. Meanwhile, the size/sound pseudoword task results correlate with those from both other tasks. This is probably because the size/sound pseudoword task used linguistic referential stimuli, like the ideophone task, and used a simple big vs. small dichotomy, like the shape/sound task. The similarities and differences of the task are illustrated below:

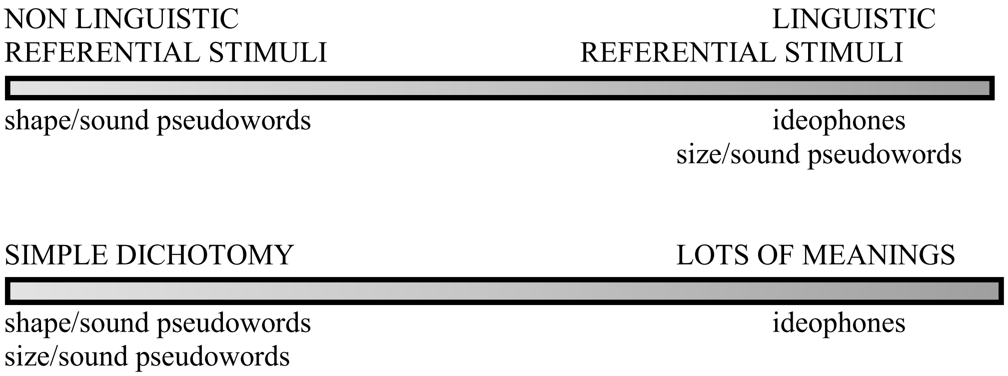


Figure 22: diagram of differences between tasks

This reaffirms the importance of looking at individual differences between participants doing different sound symbolism tasks within the same experiment, rather than considering all sound symbolism tasks to be similar.

Another interesting correlation is from within the same task. Participants who gave higher ratings to ideophones in the real condition also gave higher ratings to ideophones in the opposite condition. However, in the size/sound pseudoword task, the opposite effect happened. Participants who gave higher ratings to pseudowords in the match condition gave lower ratings to pseudowords in the mismatch condition. This is probably because the cross-modal mappings in the size/sound task are more obvious and consistent —which is unsurprising, since we deliberately designed these pseudowords to be like that— whereas the natural language ideophones are still sound-symbolic but with much more variety. Moreover, the size/sound pseudowords are symmetrical in that the match and mismatch conditions use opposite ends of the same vocalic and consonant voicing spectrum, but the ideophones do not have this perfect symmetry; the opposite translations generally don't have actual cross-modal clashes, rather, they don't have the cross-modal correspondences that the real translations do. This is borne out by the fact that the mean rating given to ideophones in the opposite condition (3.50) is higher than the mean rating given to size/sound pseudowords in the mismatch condition (2.54). We have written before that sound symbolism research needs to use real sound-symbolic words from real languages as well as deliberately-constructed pseudowords; this difference in how participants rate them is further evidence that findings from research with pseudowords should not be directly applied to real words, and vice versa.

The big question is: why was there no ERP effect?

One unlikely possibility is that ERP time window was too short, and that there might be a massive effect 1000ms-1500ms after the stimulus. However, much later effects are more common with more complicated tasks like incremental sentence processing, so it seems unlikely to be the case here.

It could also be that this rating task is too straightforward. To put it another way, perhaps sound symbolism only comes to the fore in ERPs when interacting with broader language processes. The rating effect is there behaviourally, but perhaps the simple cross-modal associations are so automatic that they are not big enough, unexpected enough, or difficult enough to show processing differences. For example, the frequency of individual letters, bigrams, and trigrams shows early ERP effects in word reading tasks or word/pseudoword judgement tasks (Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006). However, we are unaware of any studies which show an ERP effect of letter frequency in a task where participants simply read individual letters and rate how frequent they think they are. All of the previous EEG/ERP tasks with sound symbolism have more complicated and/or different tasks where participants have to do far more than simply noticing the word pairs and explicitly assessing how well they go together. In this task, participants are instructed to pay attention to the sound symbolism itself; in all the other tasks, participants have to do something else and are often unaware of the sound-symbolic manipulations.

Alternatively, there may be huge effects —as indeed there are, behaviourally— but ERP analyses are insensitive to them. This may because the effects are out of sync and average each other out, or computed over a much longer time frame, so perhaps an oscillations analysis of the data may find something. Indeed, Asano et al. (2015) found an increase in power in the gamma band in response to matching over mismatching stimuli. A time-frequency reanalysis of this dataset may find the same effect.

In conclusion, these behavioural findings confirm previous work. While this study may superficially seem to be a replication of known effects, the important new contribution is disentangling the effects of different sound symbolism tasks with the same participants. To our knowledge, this is the first time that a behavioural study has looked at the same participants doing different sound symbolism tasks. Meanwhile, the ERP null findings suggest that sound symbolism may need a more complicated task in order for ERP effects to surface. We suspect that there are many such null findings out there on various institutes' harddrives, which have never been made public because they are considered un-publishable. Sound symbolism research (and research in general) has a lot to learn from null results; all data from well-designed experiments is informative, and it would be excellent for scientific research if people made it available.

Finally, these well-designed, well-controlled, and fully rated stimulus materials should be useful to researchers wanting to take this further, and will be made available via OSF as a DIY kit for future sound symbolism experiments.

References

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex*, *63*, 196–205. http://doi.org/10.1016/j.cortex.2014.08.025

Cuskley, C., Simner, J., & Kirby, S. (2015). Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research*, 1–12. http://doi.org/10.1007/s00426-015-0709-2

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-symbolism is disrupted in dyslexia: implications for the role of cross-modal abstraction processes. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Egashira, Y., Choi, D., Motoi, M., Nishimura, T., & Watanuki, S. (2015). Differences in Event-Related Potential Responses to Japanese Onomatopoeias and Common Words. *Psychology*, *6*(13), 1653. http://doi.org/10.4236/psych.2015.613161

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30*(4), 1383–1400. http://doi.org/10.1016/j.neuroimage.2005.11.048

Jones, M., Vinson, D., Clostre, N., Zhu, A. L., Santiago, J., & Vigliocco, G. (2014). The bouba effect: sound-shape iconicity in iterated and implicit learning. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* (pp. 2459–2464). Québec.

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19–28. http://doi.org/10.1016/j.cognition.2009.08.016

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Language Sciences*, 1246. http://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016a). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016b). Synthesized Size-Sound Sound Symbolism. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society [CogSci 2016]*. Philadelphia, PA: Cognitive Science Society.

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Language Sciences*, 933. http://doi.org/10.3389/fpsyg.2015.00933

Peeters, D. (2016). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society [CogSci 2016]*. Philadelphia, PA: Cognitive Science Society.

Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de*

Psychologie Expérimentale, *65*(2), 115–124. http://doi.org/10.1037/a0022268

Sučević, J., Savić, A. M., Popović, M. B., Styles, S. J., & Ković, V. (2015). Balloons and bavoons versus spikes and shikes: ERPs reveal shared neural processes for shape–sound-meaning congruence in words, and shape–sound congruence in pseudowords. *Brain and Language*, *145–146*, 11–22. http://doi.org/10.1016/j.bandl.2015.03.011

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, *93*(1), 10–19. http://doi.org/10.1016/j.bandl.2004.07.006

Chapter 8

Discussion

Sound symbolism is real; that much is clear. Linguistics has long been influenced by de Saussure's (1916) and Hockett's (1959, 1960) writings on arbitrariness, meaning that sound symbolism was ignored until relatively recently. This thesis is one of many pieces of research to show that sound symbolism exists in addition and in complement to arbitrariness in language. Specifically, this thesis examined how sound symbolism in real language affects sentence processing and novel word learning, and how this compares to sound symbolism in pseudowords. I will first summarise the findings of each chapter and how they relate to each other, and then address some of the general themes raised throughout the thesis.

Summary

Chapter 2 gave an overview of the experimental research into sound symbolism so far. Various different studies have shown that consistent sound-symbolic associations between certain sounds and certain sensory meanings exist along similar, overlapping spectra; the vowel spectrum between the front high [i] and the mid/low back [a] and [o] maps onto lightness/heaviness, onto brightness/darkness, fast/slow speed, and small/large size, while the binary distinction of consonant voicing maps onto extreme ends of sensory meaning distinctions, such as small/large size, fast/slow speed, and light/heavy weight. Most of this research has used pseudowords, which was fine for establishing and confirming the existence of a trend in the first place, it has been clear that people associate small vowels with small size and large vowels with large size since at least the 1920s (Sapir, 1929). Designing pseudoword stimuli based on existing sound-symbolic associations to investigate sound symbolism means that the research becomes somewhat circular. Researchers must move on from two-alternative forced choice *bouba/kiki* pseudoword studies to either using real words from real languages or using better-designed pseudowords to design and test specific hypotheses. Preferably both. Some of the research reviewed in Chapter 2 has used ideophones or other sound-symbolic real words, but not many studies have. The research in this thesis was designed to address that.

Chapter 3 (published as Lockwood & Tuomainen, 2015) shows that Japanese ideophones are processed differently from arbitrary Japanese words in sentence reading. I investigated the processing of Japanese ideophones in comparison to regular arbitrary Japanese adjectives or adverbs in a sentence reading task. Japanese native speakers read a whole sentence word by word and had to judge whether the sentence made sense or not, such as in *Hanako speaks French fluentl*y vs. **Hanako cooks French fluently*. The actual experimental manipulation was about whether the Japanese word for fluently was an ideophone (*perapera*) or not (*ryuuchouni*), which participants were not aware of. ERP analyses showed an increased P2 component and increased late positive complex to the ideophones. This was the first time that the P2 component had been implicated in sound symbolism research, which means that any inferences about the mechanism behind it are exploratory. Previous research has identified the P2 as being linked to phonological processing (Dien, 2009), multisensory integration (Bien, ten Oever, Goebel, & Sack, 2012), and synaesthetic experience (Brang, Kanai, Ramachandran, & Coulson, 2010). It is possible that the P2 in this context reflects the integration of the sound of the words and the sensory information depicted by them; the P2 in response to the ideophones is greater because the greater cross-modal correspondence between the sounds of the ideophones and their sensory meanings makes it easier to integrate the information into one experience than with arbitrary words. As for the late positive complex, it may reflect an increased processing cost for ideophones in comparison to arbitrary words; there may be a trade-off between an intense depiction of events with ideophones and more easily processed description of events with arbitrary words. Where ideophones are expressive, arbitrary words are efficient; perhaps the increased late positive complex reflects the harder integration of ideophones into the sentence. An alternative account is that the LPC reflects how information is retrieved from long-term memory; a larger LPC may reflect a word being more easily retrieved (Rugg & Curran, 2007), more deeply encoded (Rugg et al., 1998; Schott, Richardson-Klavehn, Heinze, & Düzel, 2002), and/or encoded in a more imagistic way (Kelly, McDevitt, & Esch, 2009; Klaver et al., 2005).

While the exact nature of the mechanism driving these processing differences is unclear, this study makes two novel contributions to the field:

1. Using real sound-symbolic words in a whole contextual sentence; previous research has presented either pseudowords or real sound-symbolic words individually with no wider context other than an image to learn or match
2. Comparing real sound-symbolic words to real arbitrary words; previous research has only compared either congruent and incongruent pseudowords, or real sound-symbolic words in congruent and incongruent conditions.

Chapter 3 is a useful first step for experimental research into how ideophones are processed in a wider linguistic context.

Chapters 4 and 5 investigated how Dutch speakers with no knowledge of Japanese learned Japanese ideophones and arbitrary adjectives. This allowed me to test individual ideophones without the constraining sentence context and to test participants without pre-existing associations and experiences of the ideophones.

In Chapter 4 (published as Lockwood, Dingemanse, & Hagoort, 2016), participants learned Japanese words with either their real Dutch translation or their opposite Dutch translation. One group of participants learned Japanese ideophones, a second group learned arbitrary Japanese adjectives. The participants who learned the ideophones learned the ones with their real translations far better than the ones with their opposite translations; participants scored 86.1% with the real ideophone translations, but only 71.1% with the opposite ideophone translations. However, with the arbitrary adjectives, it made no difference; participants scored 79.1% with the real adjectives, and 77% with the opposite adjectives. This suggested that even if you don't speak a language with ideophones in it, whatever distinguishes ideophones from arbitrary words is recognisable and exploitable for language learning. We hypothesise that people can identify the cross-modal correspondences between the sounds of the ideophones and the sensory meanings and exploit these in order to boost word learning. After the learning and test rounds, there was also a two-alternative forced choice task, where participants were told of the manipulation and had to guess what the real translation was. Participants guessed the ideophones at 72.3% accuracy and the arbitrary words at 63% accuracy. The above-chance accuracy with the arbitrary words is due to two things. Firstly, we selected the arbitrary words which Dutch participants found easiest to guess in a two-alternative forced choice pre-test in order to be consistent with the ideophones. The fact that we selected the 38 most easily guessed arbitrary words and still found no learning effects across conditions strengthens the case for sound symbolism in ideophones. Secondly, there are likely to be trace elements of sound symbolism and expressive prosody in the recordings from the native speaker which helped participants make their decisions.

In Chapter 5 (published as Lockwood, Hagoort, & Dingemanse, 2016), I re-ran the experiment from Chapter 4 while also measuring participants' EEG. I replicated the behavioural results of Chapter 4 almost exactly — participants learned the ideophones with their real translations at 86.7% accuracy and the ideophones with their opposite translations at 71.3% accuracy, while guessing the real translations in the 2AFC task at 73% accuracy — which provides stronger evidence that the effect is real. ERP analyses showed that ideophones in the real condition elicited a much higher P3 component and late positive complex during the test round. This could have been a simple learning effect, since the P3 is often linked to learning and memory, but individual difference measures showed that there was no relation between how well participants did in a condition and how high the P3 amplitude was in that condition. However, there was a correlation between how sensitive participants were to sound symbolism (as measured by their 2AFC task accuracy) and how big the ERP amplitude difference between conditions was. This suggested that the effect here was indeed about sound symbolism, with people who were more sensitive to sound symbolism showing greater ERP differences. Moreover, this effect seemed to depend on the processing of the ideophones in the opposite condition; the people who were more sensitive to sound symbolism had a lower P3 amplitude in the opposite condition. This suggests that people who are more sensitive to sound symbolism may have more difficulty in suppressing conflicting cross-modal information. The results from Chapters 4 and 5 indicate that cross-modal correspondences between sound and meaning facilitate word learning in general, while cross-modal clashes make word learning harder for people who are more sensitive to sound symbolism.

In Chapter 6, I moved onto looking at pseudowords, which allowed me to set up a learning paradigm with three contrasts rather than two. I created a set of more specific synthesised size/sound sound-symbolic stimuli so that I could investigate the real vs. opposite translation learning effect from Chapters 4 and 5 in a more graded way, with matching, neutral (i.e. neither matching more mismatching), and mismatching conditions. Some previous research (Aveyard, 2012; Thompson & Estes, 2011) had suggested that sound symbolism might work in a graded way, where participants would learn pseudowords in a matching condition better than pseudowords in a neutral condition, and pseudowords in a neutral condition better than pseudowords in a mismatching condition. Other research (Nygaard, Cook, & Namy, 2009) had suggested that both matching and mismatching conditions would be better than a neutral condition; cross-modal information, whether correspondence or clash, may be a better scaffold than no cross-modal information whatsoever. This set up could specifically test those theories. Participants rated the stimuli in a graded way, but learned and recalled the stimuli in a way that suggested that the main behavioural effect of sound symbolism was a match boost, scaffolded by congruent cross-modal mappings. Participants learned matching pseudowords better than neutral pseudowords, but there was no difference between how well participants learned neutral and mismatching pseudowords. However, participants showed a variety of different responses; 20/60 participants showed a graded learning response, 16/60 participants showed a response where they performed worst in the neutral condition, 10/60 participants performed best in the match condition and equally in the neutral and mismatch conditions, and the other 14 were different still. This averages out as a match boost performance, but also illustrates the need for more individual differences analyses in sound symbolism experiments.

In Chapter 7, I investigated individual differences in participants across a variety of sound symbolism rating tasks. Not only has previous sound symbolism research not explored individual difference measures, to the best of my knowledge, all previous sound symbolism research papers have also only involved participants doing one particular sound-symbolic task, or multiple sets of participants doing different tasks. It is important to reiterate that there's nothing wrong with this, because it builds up a pixel-by-pixel image of sound symbolism in general. But what if sound symbolism isn't a homogenous thing? What if participants do different sound-symbolic tasks differently, and their performance in one is different from their performance in others? So, in this chapter, participants rated different sets of sound-symbolic stimuli — the ideophones and their real and opposite translations from Chapters 4 and 5, the synthesised size/sound pseudowords from Chapter 6, and some shape/sound stimuli adapted from Drijvers et al. (2015). The overall pattern of behavioural ratings was as expected: participants rated the ideophones and their real translations higher than the ideophones and their opposite translations; participants rated the size/sound pseudowords in a graded way with the match condition getting higher ratings than the neutral condition and the neutral condition getting higher ratings than the mismatch condition; and participants rated the matching shape/sound picture and pseudoword pairs higher than the mismatching pairs. However, participants performed differently across tasks. There were some correlations between the ideophone and size/sound tasks, with participants who gave low ratings to the opposite ideophones also giving low ratings to the mismatching pseudowords, and the larger the difference between a participant's real and opposite ideophone ratings, the larger the difference between that participant's match and mismatch size/sound pseudoword ratings. There were also some correlations between the size/sound and shape/sound tasks, with participants who gave high ratings to the matching size/sound pseudowords also giving high ratings to the fully matching shape/sound pseudowords, with participants who gave low ratings to the mismatching size/sound pseudowords also giving low ratings to the fully mismatching shape/sound pseudowords, and the larger the difference between a participant's match and mismatch ratings in the size/sound task, the larger the difference between that participant's full match and full mismatch ratings in the shape/sound task. However, while the ideophone and the size/sound tasks correlated, and while the size/sound and the shape/sound tasks correlated, there was no correlation between the ideophone and the shape/sound tasks. This is most likely due to two separate features of the tasks — the nature of the referential stimuli and the breadth of the semantic field of the stimuli. The referential stimuli were linguistic in both the ideophone and size/sound tasks, but non-linguistic pictures in the shape/sound task. The semantic field was wide in the ideophone tasks, with all kinds of different meanings, but narrow in both the size/sound and shape/sound tasks, with a simple dichotomy of big/small and round/spiky. This may explain why the ideophone task ratings correlated with the size/sound ratings, and why the size/sound ratings correlated with the shape/sound ratings, but why there were no correlations between the ideophone and shape/sound tasks. This chapter provides further evidence of the importance of individual difference analyses and a cautionary tale against lumping all kinds of sound symbolism together under the same umbrella. I also recorded participants' EEG for an ERP analysis. It's hard to say exactly why there were no ERP effects — there could be various reasons, such as the effect taking place over a longer time period, or the task simply being too straightforward — and this may also indicate a file drawer problem. All published ERP sound symbolism research papers involve a highly involved task where participants are often unaware of the manipulation; perhaps other researchers have tried a simple rating task and found no ERP condition differences, but have never published or openly discussed their results.

General discussion

The work presented in this PhD thesis raises several general discussion points about sound symbolism.

Ways of moving beyond 2AFC tasks

Two-alternative forced choice tasks have been really useful for establishing that sound symbolism exists in the first place. When I'm talking about my work with people who are sound symbolism novices or sceptics, the most effective way of illustrating it is asking whether *bukubuku* means *fat* or *thin*. When they almost inevitably say fat, I tell them that about 90% of other people say *fat* too. This is the gateway into discussing the more complicated aspects of sound symbolism.

The problem is that too much research relies on 2AFC tasks alone, and relying on 2AFC tasks alone creates artefactual impressions about the relationship between the sound-symbolic (psuedo)words and what they describe.

A perfect example that came up while writing this section was an impromptu cake break. A colleague was leaving the institute, and brought in two cakes to celebrate/commiserate. These cakes were well-controlled, both bought from the same shop for approximately the same price, both circular, both cut into eight slices of similar size. The only difference was that one was a cheesecake, and one was an apple tart. When people took a slice on a first-come first-served basis, everybody chose the cheesecake, and the colleague who was leaving remarked "oh, if I knew people hated apple tart so much, I'd have just brought two cheesecakes".

This real life example of a two-alternative forced choice task is perfectly useful for showing that people find cheesecake tastier than apple tart. What it doesn't show is that people find apple tart more disgusting than cheesecake, but people frequently infer from 2AFC tasks that if a tendency for X over Y in one direction holds, then a tendency for Y over X in the other direction must also hold. To compare, take Eddie Izzard's "cake or death?" sketch (Izzard, 1998). Izzard raises the two-alternative forced choice of cake or death, people would prefer cake: "Cake or death?" That's a pretty easy question. Anyone could answer that. Cake or death? Eh, cake please." 100% of participants chose cheesecake over apple tart. 100% of participants would almost definitely choose cake over death. This shows the appeal of cheesecake, and of cake in general, but it does not mean that apple tart and death are equivalent, despite nobody choosing either.

This is what has happened in sound symbolism research for a long time. The fact that people associate certain sounds (e.g. [b] and [o]) with round shapes over pointy shapes and other sounds (e.g. [k] and [i]) with pointy shapes over round shapes does not mean that the sounds [k] and [i] are as pointy as the sounds [b] and [o] are round. Recent research has shown that the round-pointy spectrum is mostly driven by the association between round shapes and sounds involving lip rounding creating an artefactual association between pointy shapes and pointy sounds in two-alternative forced choice tasks (Jones et al., 2014).

In this thesis, I have used three different strategies for investigating sound symbolism beyond 2AFC tasks. At the single word level, I used learning tasks with yes/no responses in chapters 4, 5, and 6, and 1-7 rating tasks in chapter 7. I also used an unrelated sentence judgement task in whole sentences in chapter 3 to compare ideophones and arbitrary words during sentence processing. Using real words in full sentences relies on having a language with enough sound-symbolic and arbitrary words to test, which seems to limit investigations to whatever associations already exist in the language. However, it could easily be adapted for more tightly-controlled pseudowords. Let's say you want to investigate whether associations in the speed domain are graded (like in the size domain) or dominated by one association (like roundness in the round-spiky domain). Instead of using 2AFC tasks by presenting the word fast and making participants choose between the pseudowords *zizi* and *fofo* (or presenting the pseudoword *zizi* and making participants choose between fast and slow), you could set it up in a full sentence. For a related task, you could measure ERPs during sentence reading and after the sentence ask the participants to rate how well the pseudoword conveyed the intended meaning which would be obvious from the context. An unrelated task would be to choose whether the sentence was sensible and measure ERPs like in chapter 3 of this thesis. All this can be done with sentences like the following:

Iconic?	Sensible?	Sentence
yes	yes	"We arrived on time because the train went really *zizi* throughout the journey"
no	yes	"We arrived late because the train went really *zizi* throughout the journey"
yes	no	"We arrived on time because the train went really *zizi* throughout the dog"
no	no	"We arrived late because the train went really *zizi* throughout the dog"
no	yes	"We arrived on time because the train went really *fofo* throughout the journey"
yes	yes	"We arrived late because the train went really *fofo* throughout the journey"
no	no	"We arrived on time because the train went really *fofo* throughout the dog"
yes	no	"We arrived late because the train went really *fofo* throughout the dog"

Table 1: example stimuli sentences for whole sentence pseudoword ERP experiments

This approach moves beyond 2AFC tasks and would be far more informative by investigating cross-modal associations in tasks much closer to natural language use. To do so, though, requires less focus on which sounds are associated with which meanings, and more focus on how sound symbolism works in real language.

Real words and pseudowords

Sound symbolism research has mostly been conducted with deliberately-constructed pseudowords which exaggerate the sound-symbolic associations that the researcher is investigating. At the beginning of this thesis, I wrote that real sound-symbolic words, such as Japanese ideophones or Dutch onomatopoeia (Peeters, 2016), should be used instead of pseudowords. This was a rather strong version of my ecological validity drive in iconicity research, and I did initially feel vindicated from the strong behavioural and ERP effects in chapters 4 and 5. However, I realised that using real words couldn't fully tease out the differences between cross-modal correspondences, cross-modal clashes, and a cross-modally neither one thing nor the other condition, which was necessary to see whether the learning effect was a match boost effect, a graded effect, or a mismatch difficulty effect.

It is perfectly fine to use pseudowords if they are used in a way that addresses a specific question or theory that real sound-symbolic words cannot answer. In chapter 6, for example, we used pseudowords in order to add a neutral/neither condition between real and opposite translations, because that could provide extra information about whether the sound-symbolic bootstrapping effect was based on a match boost, a mismatch hindrance, both, or neither. Pseudowords are also useful for investigating as many sound associations as possible. Real ideophones won't use every single sound available in a given language in an evenly distributed way; it would be impossible to investigate, say, the effect of consonant voicing on size associations in a properly controlled way if the ideophones in that language have 80% voiceless consonants and 20% voiced consonants, and if labial consonants are really common but velar consonants are rare. With pseudowords, you can create balanced and fair experimental stimuli to address specific questions or theories, but unfortunately this is often not the case.

Moreover, the pseudowords that are used are often poorly controlled for prosody — they are often recorded by native speakers who are aware of the task, and who may be unknowingly making subtle prosodic differences. They are also often poorly controlled for composition — *bouba* and *kiki* are imbalanced in orthographic length, vowel length, and vowel identity, and yet 2AFC tasks in shape/sound sound symbolism research are widely referred to as "*bouba/kiki* tasks".

Using real sound-symbolic words should still be the default in sound symbolism research. They offer a vast set of advantages for exploring iconicity in language. For sound symbolism research that aims to investigate which associations there are between certain sounds and certain meanings, real words offer a set of sensory sound-meaning associations which have developed naturally over time and which cover a broader set of meanings than what *bouba/kiki*-style pseudowords offer (Dingemanse, 2012; Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Nygaard et al., 2009). There have not been many corpus studies of languages with extensive sound symbolism; the ones that exist are either promising beginnings (Kwon & Round, 2014; Otis & Sagi, 2008) or in need of reappraisal (Hamano, 1986, 1998). For sound symbolism research that aims to investigate the origins of language and whether iconic vocal communication enabled fully developed language, real sound-symbolic words provide a case study in how sound symbolism affects learning and development, both as a native language with infants (Imai, Kita, Nagumo, & Okada, 2008; Kantartzis, Imai, & Kita, 2011; Laing, 2014; Yoshida, 2012) and as an additional language with adults (Lockwood, Dingemanse, et al., 2016; Lockwood, Hagoort, et al., 2016; Nygaard et al., 2009). For sound symbolism research that aims to investigate how iconicity affects communication in conversation, real sound-symbolic words are perfect for showing how people use sound symbolism in a special way, often with particular gestures and different prosody (Dingemanse, 2013; Dingemanse & Akita, in press).

The insights into real language that real sound-symbolic words provide are unparalleled in pseudowords, and yet ideophones remain underused in experimental research. This perhaps illustrates the difference between language as known by linguists and language as known by psychologists. It is definitely a challenge for psychologists to adapt materials documented by linguists, which is why I hope that this thesis will inspire other researchers to use the same stimuli. The Japanese stimuli used throughout this thesis are openly available online on the Open Science Framework, as are around 200 ideophones and translations from five different languages from a separate study by Dingemanse et al. (2016). There are many rating tasks, memory tasks, and sentence tasks which can be done with these ideophones. One logical extension is to take the pseudoword sentence tasks outlined in the previous section. The task would be an unrelated sentence sense judgement task with ideophones in congruent and incongruent contextual meanings, perhaps as follows:

Iconic?	Sensible?	Sentence
yes	yes	"The puppy really *fuwafuwa* to the touch when I stroked it"
no	yes	"The snail felt really *fuwafuwa* to the touch when I stroked it"
yes	no	"The puppy really *fuwafuwa* to the touch when I looked at it"
no	no	"The snail felt really *fuwafuwa* to the touch when I looked at it"

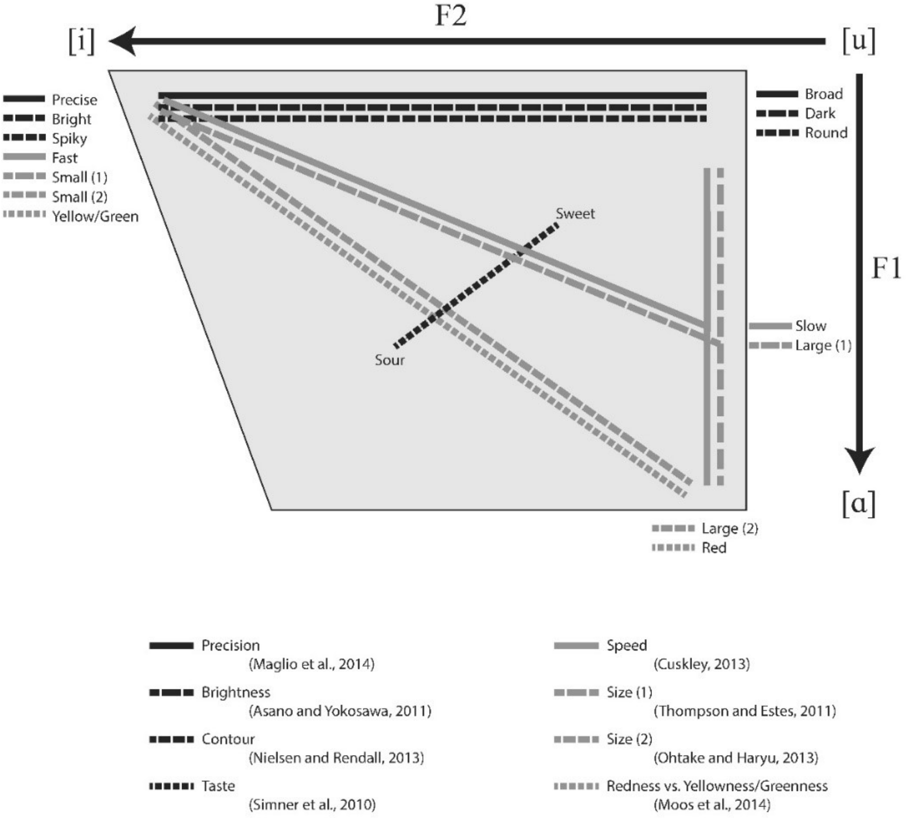
Table 2: example stimuli sentences for whole sentence ideophone rating experiments

If this thesis can inspire sound symbolism researchers to use ideophones from other languages in various experimental contexts, I will consider it a success. There is still scope for using pseudowords when they address specific questions that real words can't answer, but far, far more sound-symbolism research with real words is needed.

Sound symbolism: a heterogeneous thing

One thing that became clear throughout this thesis is that sound symbolism is an umbrella term. It covers various relationships between form and meaning, but more pertinently, it covers various *types* of relationships between form and meaning, and we must be careful not to automatically apply what happens in one sound-symbolic domain to another.

Sound symbolism research has shown that there are cross-modal correspondences between sounds and sensory meanings which people can recognise and exploit. This applies to various domains; this was covered in depth in the literature review in chapter 2, and illustrated in the diagram of vowel associations below.



Lockwood & Dingemanse (2015) *Frontiers in Psychology* — doi.org/10.3389/fpsyg.2015.01246

Figure 1: vowel space sound-symbolic associations

Beyond detailing the specific associations between certain sounds and certain meanings, there has been a tendency to group all types of sound symbolism together as one, and assume that research in one (such as pseudowords in a shape/sound task) can inform interpretations of another (such as ideophones). In fact, different types of sound symbolism work differently, as shown in chapter 7 of this thesis. However, this has been obscured by two things: firstly, there has been very little research where the same participants did different sound symbolism tasks, and secondly, 2AFC tasks have created artefactual dichotomies which have been overgeneralised to all types of sound symbolism.

There were already suggestions from research in linguistics that interpretations of sound symbolism may not be generalisable to different domains. Akita (2013) writes about the Lexical Iconicity Hierarchy for sound-symbolic words across languages. The hierarchy sets out how iconic certain sound-symbolic words are; animal imitations are more iconic than phenomimes (sound-symbolic words referring to sounds in the real world), phenomimes are more iconic than psychomimes (sound-symbolic words depicting internal psychological states), and psychomimes are more iconic than regular arbitrary words. The more iconic ideophones tend to be realised in the periphery of a clause, while the less iconic ideophones tend to be integrated further into the centre of the clause, suggesting that not only is there a hierarchy of how iconic certain domains are, this is reflected in how the language integrates them into the grammar. Furthermore, Dingemanse identified an implicational hierarchy of ideophone systems about the kinds of sound-symbolic words a language has (2012). If a language has psychomimes, the least iconic sound-symbolic word type, it also has all the others; but if a language has animal imitations, it does not necessarily also have psychomimes or phenomimes. That words differ in how sound-symbolic they are in different domains, and that this hierarchy of iconicity is borne out in which domains languages have sound-symbolic words for, is a clear indicator that not all types of sound symbolism work in the same way. This is further suggested by Dingemanse et al.'s (2016) research which showed that Dutch participants are not equally sensitive to ideophones from different categories from different languages; ideophones expressing sound were most accurately guessed, while ideophones expressing texture were least accurately guessed (though still above chance).

Chapter 7 of this thesis looked at how the same participants would perform the same task across three different sets of sound-symbolic stimuli; Japanese ideophones, size/sound pseudowords, and shape/sound pseudowords. This allowed us to investigate whether different sound-symbolic stimuli are rated similarly or not, as well as allowing us to explore individual differences (more on that in the next section). There were differences between all three tasks. The size/sound pseudowords had the most sound-symbolic saliency, with the highest ratings for matching stimuli and the lowest ratings for mismatching stimuli and very little variation within ratings. There was a very obviously graded effect to size/sound ratings. The shape/sound pseudowords had far less variety, although the effect was still there, suggesting that the strong *bouba/kiki* effects found in 2AFC tasks are real but may be weaker than previously thought. Moreover, the analysis found that the effect was stronger for round shapes than spiky shapes, providing additional support for the idea that spikiness-roundness is not a graded sound-symbolic continuum. Meanwhile, the ideophones task showed that participants rated the opposite translations quite neutrally, despite being worse at learning them in chapters 4 and 5. There was also a lot of variation in how the ideophones were rated.

While there were some correlations between how participants rated the ideophones and the size/sound stimuli, and some correlations between how participants rated the size/sound stimuli and the shape/sound stimuli, there were clear differences in how the same participants treated the same tasks with different stimuli.

This goes to show that sound symbolism is not homogenous, and should not be treated as such. Words and pseudowords with cross-modal correspondences between sound and meaning are all examples of sound symbolism, but the way in which they are treated differs according to the nature and the sensory domain of the stimuli. The size/sound associations come out the most clearly, probably because the perceptuomotor associations between the size meanings and the size of the vocal tract are the most obvious. The space between the tongue and the roof of the mouth is big during the vowels [a] and [o], small during the vowels [i] and [ʏ], and in between the two for the vowels [ɛ] and [ə]. The shape/sound analogies are less clear; the lip rounding involved in the vowel [o] and labial consonants created the stronger likelihood for participants to rate round shapes with the "round" sounds, but there is no obvious analogy between spiky shapes and sounds like [k], [i], or other sounds which don't involve lip rounding. Even then, participants did not rate the full match shape/sound pairs as highly as the match size/sound pairs, which suggests that shape analogies are less salient than size analogies. Considering that shape/sound sound symbolism tends to be the example which is used to explain what sound symbolism is, it is important to point out that it is less salient than it is given credit for. As for the ideophone task, the 38 ideophones included a variety of different meanings from different domains, meaning that there was no one consistent perceptuomotor analogy. It may also be the case that different participants find different perceptuomotor analogies more salient in ideophones; one participant may easily identify the analogy between the light frication and approximation of the consonants in *fuwafuwa* and the texture of the fluffy texture it describes, but miss the analogy between the lip rounding in bukubuku and the roundness of the fat object it describes, while this may be vice versa for another participant.

Future sound symbolism research in all domains is needed, but it is important not to draw far-reaching conclusions about sound symbolism as a whole from experiments on a single type of sound symbolism, and about sound symbolism in natural language from experiments using pseudowords.

Individual differences in sound symbolism

It's not just sound symbolism itself that is not homogenous; neither is sensitivity to sound symbolism across different individuals. This is only to be expected — even the purported figure of 95% of people choosing *bouba* with a round figure and *kiki* with a spiky figure is presumably based on a mean score (Ramachandran & Hubbard, 2001) — but the vast majority of sound symbolism research to date has focused on group effects rather than considering whether, why, and how much results may vary across participants.

Whole group effects have been useful in what is a new-ish and small-ish field. Saying that 95% of people choose cross-modally congruent stimuli, or that people learn ideophones better with the real translations than their opposite translations by 86.1% to 71.1%, has been hugely important for showing people that sound symbolism exists and that sound symbolism works for most people. However, this only takes the field so far, and sound symbolism research should investigate individual differences more thoroughly, because the "variation is the essential data, the levers that give one insight into how mental processes work" (Levinson, 2012).

A case in point is chapter 5. The behavioural results show that, on average, people get 86.7% in the real condition and 71.3% in the opposite condition, which could be summed up in a simple histogram with 95% confidence intervals. But simply plotting dot plots of the data shows that this effect is actually consistent across participants, regardless of how well they did overall; 23 of the 29 participants were better in the real condition than in the opposite condition, and of the six who didn't, two performed equally in both conditions, and the other four did only slightly better in the opposite condition with a mean 3.95% percentage point difference. The ERP results show that the P3 and late positive complex are higher in response to the ideophones in the real condition than the ideophones in the opposite condition. But correlating participants' ERP amplitude difference with their 2AFC task performance to gauge sound-symbolic sensitivity, and plotting separate ERP graphs for the top half and bottom half of participants according to their sound-symbolic sensitivity, was highly informative. This showed that the ERP effect was huge for people who were shown to be more sensitive to sound symbolism in an independent task, and barely there at all for people who were less sensitive to sound symbolism. Moreover, it showed that this was driven by the response to ideophones in the opposite condition, not ideophones in the real condition.

Without looking at individual differences, the variation, and the insights into what contributes to it, would have been lost. Instead, it showed that the vast majority of people can use cross-modal correspondences to bootstrap word learning, but that around half those people are able to easily ignore cross-modal clashes while the other half have to try to ignore that conflicting information. This is only based on one study so far; some of these differences may simply be noise and more research is needed before being relatively certain, but I speculate that individual variation in sound-symbolic sensitivity may look like this:

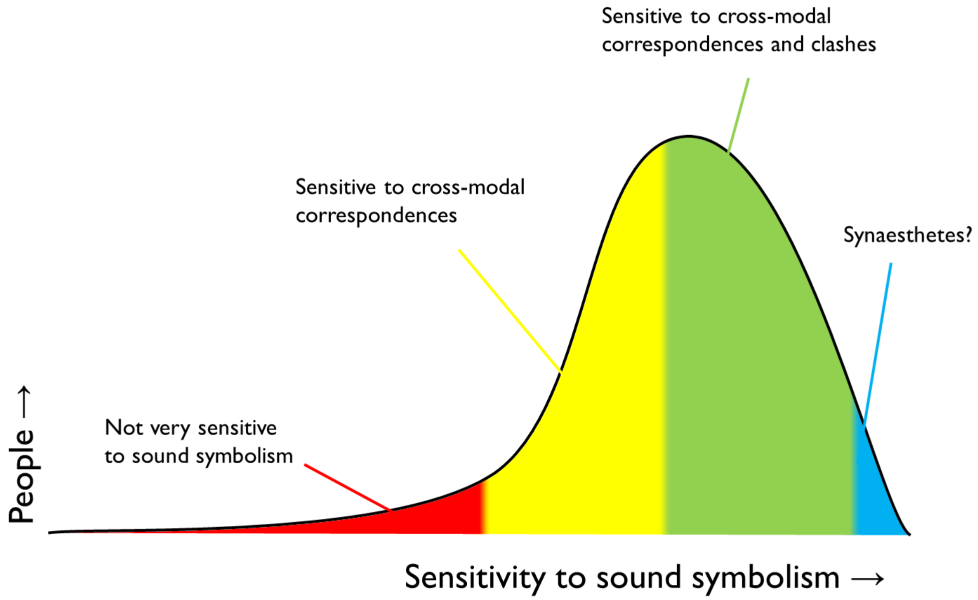


Figure 2: sketch of how sensitivity to sound symbolism may vary across the population

This is only a preliminary description which needs corroboration from further experiments. Indeed, part of the Groot Nationaal Onderzoek project was designed to look at how synaesthesia interacts with sound symbolism, and our preliminary findings suggest that synaesthetes are better at the Japanese ideophone 2AFC task when controlling for age, gender, education level, and whether or not participants are dyslexic according to self-report (Lockwood, van Leeuwen, Drijvers, & Dingemanse, 2016). Further experiments could also take the opposite direction; gauge participants' sound-symbolic sensitivity from a 2AFC task, and then see how that affects other cross-modal processing tasks. For example, participants could do a pitch/size task where they have to learn items with congruent (low pitch, large size / high pitch, small size) and incongruent (low pitch, small size / high pitch, large size) mappings. Based on the findings in this thesis, all participants would show a congruence boost, while participants who are more sensitive to sound symbolism may find it harder to suppress the incongruent mappings. A further point that needs investigating is whether people are differently sensitive to different types of sound symbolism. For example, somebody could rate the size/sound stimuli from chapter 7 to the extremes, consistently choosing 7 in the match condition and 1 in the mismatch condition, but still only get around 50% in the 2AFC task for guessing the meanings of Japanese ideophones from chapters 4 and 5. I haven't tested that in this thesis, but given the heterogeneity of sound symbolism outlined earlier, it's certainly possible.

While showing that sound symbolism is real with group effects has been instrumental in the development of the field, the next exciting area for sound symbolism research is to investigate individual differences for a fuller explanation of how sound symbolism works.

Possible mechanisms

So much for what sound symbolism is (and isn't) and the variation across different people. How does it work?

Throughout this thesis, I have referred to sound symbolism as cross-modal correspondences between the sound of a word and its sensory meaning. In order to explore how sound symbolism works, it is firstly important to discuss what cross-modal correspondences are.

Cross-modal correspondences happen when a dimension of one sensory modality (e.g. high pitch) is associated with a dimension of another sensory modality (e.g. small size). One school of thought is that cross-modal correspondences are a weak form of synaesthesia (Marks, 1978; Martino & Marks, 2001; Ward, Huckstep, & Tsakanikos, 2006), whereby all people make connections between various different sensory modalities, and synaesthetes have an extreme, exaggerated, and idiosyncratic version of it. The opposite view is that cross-modal correspondences are entirely independent of synaesthesia and are mostly learned from the environment, but the conflation of the two happens because of their shared superficial features (Parise & Spence, 2012; Spence, 2011).

Both have some merit to them, and as with many polemics, the true nature is probably something of a combination. The fact that both synaesthetes and non-synaesthetes make similar grapheme-colour associations suggests a greater deal of shared processing than a strictly learned model could account for (Moos, Smith, Miller, & Simmons, 2014; Simner et al., 2005); that synaesthetes' associations correlate with the colours of children's fridge magnets suggests that some of it is learned from, or at least reinforced by, the environment (Witthoft & Winawer, 2006; Witthoft, Winawer, & Eagleman, 2015).

In terms of sound symbolism in language, the results presented in this thesis fall closer to the side of weak synaesthesia than learned associations; the results suggest that the vast majority of people are sensitive to cross-modal correspondences in language, and that this is unlikely to have been learned independently of a multisensory mechanism which underpins several aspects of both synaesthesia and cross-modal correspondences. Before exploring the possible links between sound symbolism and synaesthesia, I will first discuss how sound symbolism is unlikely to be based on learned associations and how perceptuomotor analogies form only part of the picture.

The correspondences between Dutch meanings and Japanese words are unlikely to be learned from the environment, given that none of my Dutch participants had any knowledge of Japanese, let alone Japanese ideophones. Language is also that much more abstract than direct sensory properties, so while an infant may develop cross-modal correspondences between pitch and size from working out that big things tend to make lower noises than small things, the real world learning source of bright things making [i] noises and dark things making [a] noises is less obvious. Moreover, the cross-modal correspondences between sound and meaning in Japanese are not easily documentable. Hamano (1986, 1998) deconstructed Japanese ideophones into sets of meanings attributed to each phoneme. Each phoneme brings particular meanings to an ideophone depending on its position in the ideophone; for example, vowels in the second position of reduplicated CVCV ideophones (e.g. the /a/ in *kirakira*) contribute the following properties (Hamano, 1986, p. 150):

/i/ = +tense, +small, -large, -protrusion
/a/ = -tense, -small, +large, -protrusion
/o/ = -tense, -small, -large, -protrusion
/u/ = -tense, +small, -large, +protrusion

This was an admirable effort to gather hundreds of examples of Japanese ideophones and average them out into distinct units which could be constructed à la carte, but it often results in a vague and sometimes metaphorical interpretation of each individual phoneme. For example, "The alveolar stop /t/ is used in contexts where 'hitting' is involved. The meaning is also broadened to include the meanings of 'coming into close contact' and 'complete agreement.' " (Hamano, 1986, p. 177). These deconstructed associations per phoneme are vague to the point of being potentially unlearnable from the environment. It seems highly unlikely that Dutch participants speaking no Japanese whatsoever could have learned these associations, but they may reflect synaesthetic tendencies to associate qualities or characteristics to letters and sounds.

A separate, synaesthesia-agnostic view is that cross-modal correspondences between sound and meaning are perceptuomotor analogies (Dingemanse et al., 2015). The perceptuomotor analogies for some types sound symbolism with are obvious — round lip shapes for round shapes, large and small oral cavities for large and small vowels — but the picture is mixed with ideophones. For some, the mappings are clear, such as with *bukubuku*; the roundness of the lips in the [b] and [u] sounds mirrors the roundness of the fat object depicted. In other cases, the mappings are not as immediately transparent, the possible perceptuomotor analogies become more convoluted; what analogies do Dutch participants make between *fonkelend*, or shining/glistening, and *kirakira*? Perhaps the [k] and the [i] feel bright, the forward-moving tongue while making the [r] sound feels like it represents the light emanating forward from a source, and the large [a] oral cavity represents the light spreading and disseminating over a wider area. This approach resembles Hamano's methods and reasoning, and while it may capture tendencies of how phonemes get used in Japanese ideophones, it is probably both highly subjective and implicit. When I asked my participants about any strategies they had during the Japanese ideophone learning tasks, the most common responses were thinking up convoluted links to other words or things ("*fuwafuwa* is 'fluffy' and I thought of my cat who's called Fifi and she's fluffy") or that there was no particular strategy ("it just kind of felt that way"). None of my participants suggested that they associate particular phonemes with particular sensory properties. That said, the perceptuomotor anologies don't encode the full meanings of the ideophones, they merely provide small cues towards the meanings, which is generally enough to push people closer to one interpretation than another in 2AFC tasks, and may be enough to latch onto for scaffolding word learning. If there are perceptuomotor analogies involved in sound symbolism with ideophones, they are to some extent implicit; and if they are to some extent implicit, then that suggests that there is a deeper sensory mechanism which underpins them.

Back to synaesthesia, then. The results in this thesis do not test synaesthetes directly, and cannot therefore say whether sound symbolism is a weak form of synaesthesia. However, there is certainly an overlap between synaesthesia and sensitivity to sound symbolism. Research which has tested synaesthetes has shown that synaesthetes are slightly more sensitive to sound symbolism than non-synaesthetes (Bankieris & Simner, 2015; Lacey, Martinez, McCormick, & Sathian, 2016), and our Groot Nationaal Onderzoek project also suggests that synaesthetes are more sensitive to sound symbolism (Lockwood, van Leeuwen, et al., 2016).

There are two neural theories of synaesthesia, which have been developed from experiments with grapheme-colour associations. One theory states that there is direct cross-wiring between adjacent cortical areas, meaning that extra connections drive this cross-activation. The other theory states that this cross-activation is indirect, driven by disinhibited feedback from the superior parietal lobe during multimodal integration. Van Leeuwen (2011) reconciled the theories by showing that both were true for different types of synaesthetes — projector synaesthetes, who see grapheme-induced colours out there in the real world, have cross-wiring, while associator synaesthetes, who see or feel grapheme-induced colours strongly internally in their mind's eye, have disinhibited feedback.

If the cross-modal associations found in sound symbolism are a weak form of synaesthesia, it is more likely to be associator-type synaesthesia. The link between small size and "small" vowels isn't perceived externally in the real world, it "just kind of feels that way". Indeed, there is some tentative evidence for sound symbolism leading to higher activation in the superior parietal lobe (Revill, Namy, DeFife, & Nygaard, 2014). Perhaps, then, a semi-synaesthetic account of sound symbolism could work in the following way, based on the Hickok and Poeppel (2007) model of speech and van Leeuwen et al.'s model of associator synaesthesia (2011):

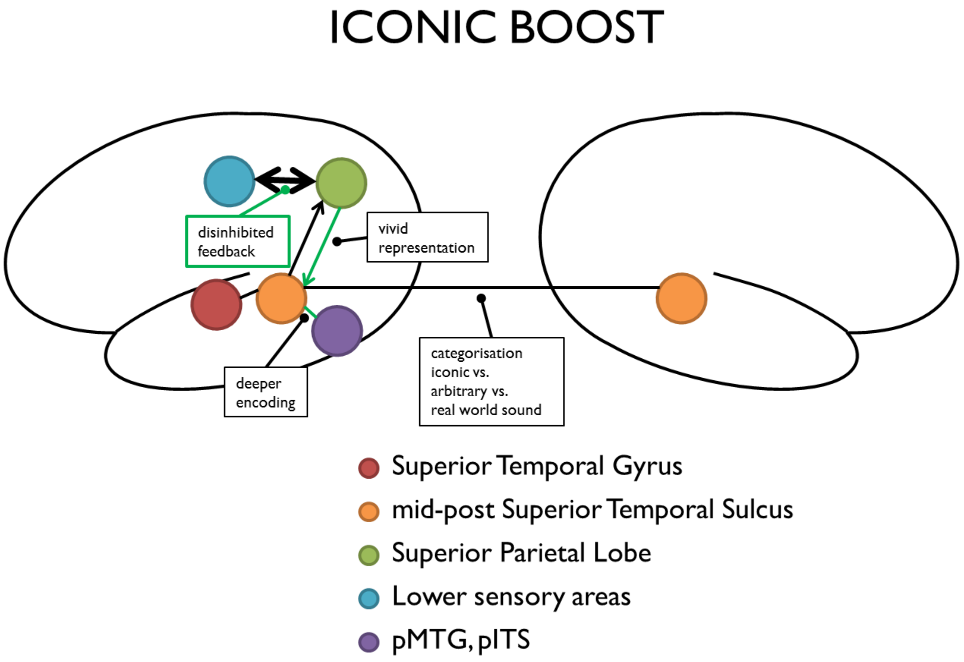


Figure 3: possible iconic word learning model

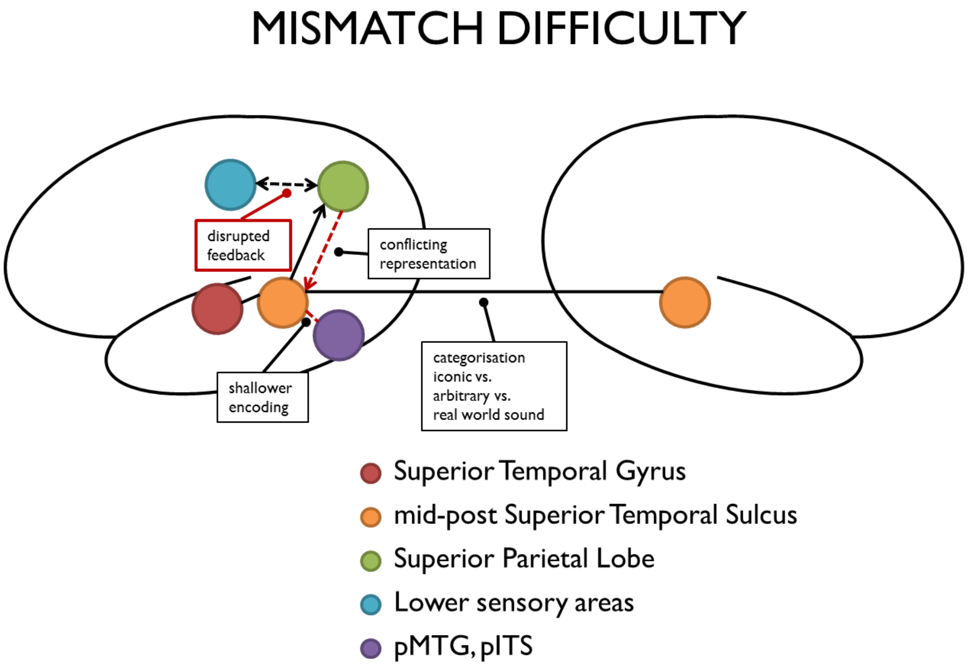


Figure 4: possible iconic mismatch word learning model

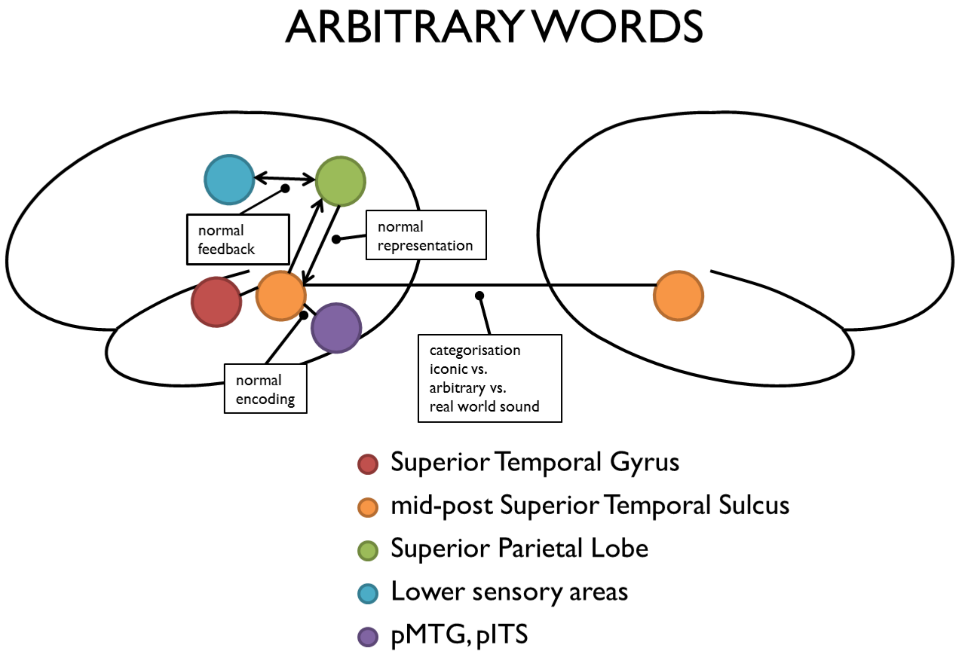


Figure 5: possible arbitrary word learning model

This sketch adapts the Hickok and Poeppel (2007) model of speech and the van Leeuwen (2011) model of associator synaesthesia. Speech perception and phonological processing takes place in the Superior Temporal Gyrus (STG) and Superior Temporal Sulcus (STS), after which information is fed through to the Superior Parietal Lobe (SPL), a known hub of multimodal integration (Bien et al., 2012; Molholm et al., 2006; Wolpert, Goodbody, & Husain, 1998). Moreover, the Superior Parietal Lobe has been heavily implicated in synaesthesia (Hubbard & Ramachandran, 2005; Rouw, Scholte, & Colizoli, 2011), with increased structural connectivity in the SPL found in synaesthetes (Rouw & Scholte, 2007) and TMS to the SPL found to disrupt synaesthetic experience (Muggleton, Tsakanikos, Walsh, & Ward, 2007; Rothen, Nyffeler, von Wartburg, Müri, & Meier, 2010). The pMTG and pITS handle the interface with lexical processing under Hickok and Poeppel’s (2007) model, and are included for completeness.

The sound symbolism imaging literature is somewhat scant, but the studies that do exist find that sound-symbolic words elicit higher activation in left SPL (Revill et al., 2014) and in the right STS (Kanero, Imai, Okuda, Okada, & Matsuda, 2014). Revill et al. argue that the activation of the left SPL shows that sound symbolism is related to a more general sensory processing mechanism, while Kanero et al. write that the activation of the right STS shows ideophones' dual nature as both linguistic item and depictive linguistic gesture.

The synaesthesia literature posits that disinhibited feedback between the SPL and lower sensory areas result in the additional sensory activation elicited by synaesthetic inducing stimuli. To apply the same mechanism to iconic stimuli, this additional sensory activation leads to a more vivid representation of sound-symbolic words, while perhaps disrupted feedback for iconic mismatches may lead to a less secure representation. The strength of the representation then in turn affects word learning; arbitrary words with regular representations are learned in a regular way, iconic words with vivid representations are learned better, mismatching words with insecure representations may or may not be learned worse depending on the participant.

Perhaps the two can be reconciled in the following way. Sound symbolism is based on cross-modal associations. These cross-modal associations come about from disinhibited feedback from the Superior Parietal Lobe, which cascades information back down to lower sensory areas. This additional sensory activation during word learning (both as a native speaker and naïve participant learning individual words) makes sound-symbolic words more vivid or imagistic. The vividness of the sound-symbolic words, coupled with the marked phonology and/or prosody that ideophones often have, allows the right Superior Temporal Sulcus to categorise the words as iconic or not. The vividness and iconic status leads to better encoding and deeper learning of sound-symbolic words in comparison to arbitrary words or words with no cross-modal correspondences.

Different stages of this sketch may be reflected in the ERP results in this thesis. In Japanese participants in Chapter 3, the P2 may reflect the initial disinhibited feedback from the cross-modal correspondences in the ideophones compared to the arbitrary words, while the late positive complex may reflect the deeper encoding of the ideophones compared to the arbitrary words in the speakers' lexicons. In Dutch participants in Chapter 5, the P3 amplitude differences may reflect the individual differences in how difficult people find it to suppress the disrupted feedback and conflicting cross-modal information during word learning, while the late positive complex may again reflect the deeper encoding of the ideophones in the real condition compared to the opposite condition.

Further research could investigate this by using the ideophone learning paradigm in chapters 4 and 5 with fMRI and dynamic causal modelling (DCM), and to do the same or similar experiments with confirmed synaesthete participants. TMS experiments similar to the ones done with synaesthetes would also be informative; this sketch would predict that TMS to the SPL would affect the differing levels of feedback that create the cross-modal associations, and would therefore reduce or remove the effect of iconicity. With Dutch participants, TMS to the SPL before the ideophone learning task should result in equal performance with ideophones and arbitrary words. With Japanese participants, TMS to the SPL before the sentence assessment task may result in no ERP effects. Revill et al. (2014) found functional anisotropy differences in the Superior Longitudinal Fasciculus according to sound symbolism task performance; DCM or DTI analyses may show increased functional or structural connectivity in the SPL in participants who are more sensitive to sound symbolism in the same way that synaesthetes have increased functional/structural connectivity compared to non-synaesthetes.

Final summary

The sounds of language and the sensory information those sounds convey are tightly, but mysteriously, intertwined. Von Humboldt (1836) wrote that "to represent outer objects that speak to all senses at once, and the inner motions of the mind, entirely by impressions on the ear, is an operation largely inexplicable in detail". But 180 years ago, EEG hadn't been invented, and Japan was an isolationist island which had closed its borders, making both brain imaging and a rich inventory of ideophones inaccessible. This thesis proposes that to represent outer objects that speak to all (or at least some) senses at once, and the inner motions of the mind, entirely by impressions on the ear, is facilitated by cross-modal correspondences between sound and sensory meaning, and these correspondences facilitate word learning and retrieval through a disinhibited feedback mechanism which may be the same as the mechanism underpinning synaesthesia.

References

Akita, K. (2013). The lexical iconicity hierarchy and its grammatical correlates. In L. Elleström, O. Fischer, & C. Ljungberg (Eds.), *Iconic Investigations* (pp. 331–349). Amsterdam: John Benjamins Publishing.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & Cognition*, *40*(1), 83–92. http://doi.org/10.3758/s13421-011-0139-3

Bankieris, K., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition*, *136*, 186–195. http://doi.org/10.1016/j.cognition.2014.11.013

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663–672. http://doi.org/10.1016/j.neuroimage.2011.06.095

Brang, D., Kanai, S., Ramachandran, V. S., & Coulson, S. (2010). Contextual Priming in Grapheme–Color Synesthetes and Yoked Controls: 400 msec in the Life of a Synesthete. *Journal of Cognitive Neuroscience*, *23*(7), 1681–1696. http://doi.org/10.1162/jocn.2010.21486

Dien, J. (2009). The neurocognitive basis of reading single words as seen through early latency ERPs: A model of converging pathways. *Biological Psychology*, *80*(1), 10–22. http://doi.org/10.1016/j.biopsycho.2008.04.013

Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, *6*(10), 654–672. http://doi.org/10.1002/lnc3.361

Dingemanse, M. (2013). Ideophones and gesture in everyday speech. *Gesture*, *13*(2), 143–165. http://doi.org/10.1075/gest.13.2.02din

Dingemanse, M., & Akita, K. (in press). An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, *19*(10), 603–615. http://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What Sound Symbolism Can and Cannot Do: Testing the Iconicity of Ideophones from Five Languages. *Language*, *92*(2), e117-33. http://doi.org/10.1353/lan.2016.0034

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-symbolism is disrupted in dyslexia: implications for the role of cross-modal abstraction processes. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Hamano, S. (1986). *The Sound-Symbolic System of Japanese*. Retrieved from http://archive.org/details/soundsymbolicsys00hama

Hamano, S. (1998). *The Sound-Symbolic System of Japanese*. Center for the Study of Language and Information ; Tokyo.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. http://doi.org/10.1038/nrn2113

Hockett, C. (1959). Animal 'languages' and human language. *Human Biology*, *31*(1), 32–39.

Hockett, C. (1960). The Origin of Speech. *Scientific American*, *203*, 89–97.

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive Mechanisms of Synesthesia. *Neuron*, *48*(3), 509–520. http://doi.org/10.1016/j.neuron.2005.10.012

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. http://doi.org/10.1016/j.cognition.2008.07.015

Izzard, E. (1998). Cake Or Death? from Dress To Kill. Retrieved 13 July 2016, from https://www.youtube.com/watch?v=PVH0gZO5lq0

Jones, M., Vinson, D., Clostre, N., Zhu, A. L., Santiago, J., & Vigliocco, G. (2014). The bouba effect: sound-shape iconicity in iterated and implicit learning. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* (pp. 2459–2464). Québec.

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How Sound Symbolism Is Processed in the Brain: A Study on Japanese Mimetic Words. *PLoS ONE*, *9*(5), e97905. http://doi.org/10.1371/journal.pone.0097905

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese Sound-Symbolism Facilitates Word Learning in English-Speaking Children. *Cognitive Science*, *35*(3), 575–586. http://doi.org/10.1111/j.1551-6709.2010.01169.x

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, *24*(2), 313–334. http://doi.org/10.1080/01690960802365567

Klaver, P., Fell, J., Dietl, T., Schür, S., Schaller, C., Elger, C. E., & Fernández, G. (2005). Word imageability affects the hippocampus in recognition memory. *Hippocampus*, *15*(6), 704–712. http://doi.org/10.1002/hipo.20081

Kwon, N., & Round, E. R. (2014). Phonaesthemes in morphological theory. *Morphology*, 1–27. http://doi.org/10.1007/s11525-014-9250-z

Lacey, S., Martinez, M., McCormick, K., & Sathian, K. (2016). Synesthesia strengthens sound-symbolic cross-modal correspondences. *European Journal of Neuroscience*, n/a-n/a. http://doi.org/10.1111/ejn.13381

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, *34*(5), 387–405. http://doi.org/10.1177/0142723714550110

Levinson, S. C. (2012). The Original Sin of Cognitive Science. *Topics in Cognitive Science*, *4*(3), 396–403. http://doi.org/10.1111/j.1756-8765.2012.01195.x

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-Symbolism Boosts Novel Word Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. http://doi.org/10.1037/xlm0000235

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How Iconicity Helps People Learn New Words: Neural Correlates and Individual Differences in Sound-Symbolic Bootstrapping. *Collabra*, *2*(1). http://doi.org/10.1525/collabra.42

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Language Sciences*, 933. http://doi.org/10.3389/fpsyg.2015.00933

Lockwood, G., van Leeuwen, T., Drijvers, L., & Dingemanse, M. (2016, 23/4). *Synaesthesia and sound symbolism - insights from the Groot Nationaal Onderzoek project (poster)*. Presented at the Synaesthesia and Cross-Modal Perception, Dublin. Retrieved from https://dx.doi.org/10.6084/m9.figshare.3184531.v1

Marks, L. E. (1978). *The unity of the senses: interrelations among the modalities*. New York; London: Academic Press.

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and Weak. *Current Directions in Psychological Science*, *10*(2), 61–65. http://doi.org/10.1111/1467-8721.00116

Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., … Foxe, J. J. (2006). Audio-Visual Multisensory Integration in Superior Parietal Lobule Revealed by Human Intracranial Recordings. *Journal of Neurophysiology*, *96*(2), 721–729. http://doi.org/10.1152/jn.00285.2006

Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: Vowel colours in the ear of the beholder. *I-Perception*, *5*(2), 132–142. http://doi.org/10.1068/i0626

Muggleton, N., Tsakanikos, E., Walsh, V., & Ward, J. (2007). Disruption of synaesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia*, *45*(7), 1582–1585. http://doi.org/10.1016/j.neuropsychologia.2006.11.021

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*(1), 181–186. http://doi.org/10.1016/j.cognition.2009.04.001

Otis, K., & Sagi, E. (2008). Phonaesthemes: A corpus-based analysis. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 65–70).

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, *220*(3–4), 319–333. http://doi.org/10.1007/s00221-012-3140-6

Peeters, D. (2016). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society [CogSci 2016]*. Austin, TX: Cognitive Science Society.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia -- A window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. *Brain and Language*, *128*(1), 18–24. http://doi.org/10.1016/j.bandl.2013.11.002

Rothen, N., Nyffeler, T., von Wartburg, R., Müri, R., & Meier, B. (2010). Parieto-occipital suppression eliminates implicit bidirectionality in grapheme-colour synaesthesia. *Neuropsychologia*, *48*(12), 3482–3487. http://doi.org/10.1016/j.neuropsychologia.2010.07.032

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*(6), 792–797. http://doi.org/10.1038/nn1906

Rouw, R., Scholte, H. S., & Colizoli, O. (2011). Brain areas involved in synaesthesia: A review. *Journal of Neuropsychology*, *5*(2), 214–242. http://doi.org/10.1111/j.1748-6653.2011.02006.x

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. http://doi.org/10.1016/j.tics.2007.04.004

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595–598. http://doi.org/10.1038/33396

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225–239. http://doi.org/10.1037/h0070931

Saussure, F. de. (1916). *Course in general linguistics (1959 edition)*. New York: Philosophical Library.

Schott, B., Richardson-Klavehn, A., Heinze, H.-J., & Düzel, E. (2002). Perceptual Priming Versus Explicit Memory: Dissociable Neural Correlates at Encoding. *Journal of Cognitive Neuroscience*, *14*(4), 578–592. http://doi.org/10.1162/08989290260045828

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. A. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology*, *22*(8), 1069–1085. http://doi.org/10.1080/02643290500200122

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. http://doi.org/10.3758/s13414-010-0073-7

Thompson, P., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *The Quarterly Journal of Experimental Psychology*, *64*(12), 2392–2404. http://doi.org/10.1080/17470218.2011.605898 van Leeuwen, T. M., den Ouden, H. E. M., & Hagoort, P. (2011). Effective connectivity determines the nature of subjective experience in grapheme-color synesthesia. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(27), 9879–9884. http://doi.org/10.1523/JNEUROSCI.0569-11.2011

von Humboldt, W. (1836). *On Language: On the Diversity of Human Language Construction and Its Influence on the Mental Development of the Human Species*. (M. Losonsky, Ed.). Cambridge: Cambridge University Press.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *42*(2), 264–280.

Witthoft, N., & Winawer, J. (2006). Synesthetic Colors Determined by Having Colored Refrigerator Magnets in Childhood. *Cortex*, *42*(2), 175–183. http://doi.org/10.1016/S0010-9452(08)70342-3

Witthoft, N., Winawer, J., & Eagleman, D. M. (2015). Prevalence of Learned Grapheme-Color Pairings in a Large Online Sample of Synesthetes. *PLoS ONE*, *10*(3), e0118996. http://doi.org/10.1371/journal.pone.0118996

Wolpert, D. M., Goodbody, S. J., & Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. *Nature Neuroscience*, *1*(6), 529–533. http://doi.org/10.1038/2245

Yoshida, H. (2012). A Cross-Linguistic Study of Sound Symbolism in Children's Verb Learning. *Journal of Cognition and Development*, *13*(2), 232–265. http://doi.org/10.1080/15248372.2011.573515

Samenvatting, Curriculum Vitae, Publications, Acknowledgements, and Max Planck Institute Series in Psycholinguistics

Samenvatting

Wanneer de klank van een woord op de één of andere manier de betekenis van dat woord reflecteert, noemen we dat klanksymboliek. In Europese talen gaat klanksymboliek meestal niet veel verder dan onomatopeën – iemand *slurpt* soep , de kat *miauwt*, de auto gaat *vroem*. Daarom zagen onderzoekers klanksymboliek lange tijd slechts als een marginaal aspect van taal.

Maar klanksymboliek gaat veel verder dan de geluidseffecten in stripboeken of dierengeluiden voor kinderen. Klanksymboliek kun je terugvinden in talen over de hele wereld. In veel talen vind je ideofonen, opvallende woorden die zintuiglijke informatie verbeelden. Het Japanse woord "kirakira" bijvoorbeeld, betekent "fonkelend". Ideofonen worden in het Nederlands of Engels het best vertaald met bijvoeglijk naamwoorden of bijwoorden, maar in talen waarin ideofonen veel gebruikt worden, spelen ze ook andere rollen. Je kunt het beste over ideofonen denken als een aparte woordcategorie. Bovendien zijn ideofonen geen alles-of-niets fenomeen: net als Europese talen hebben talen met ideofonen ook normale adjectieven, bijwoorden en werkwoorden waar de relatie tussen vorm en betekenis meer willekeurig is. Ideofonen worden gebruikt als aanvulling op, niet in plaats van, gewone woorden.

Ondanks dat ideofonen perfecte voorbeelden zijn van klanksymboliek in natuurlijke taal, zijn ze nog niet vaak gebruikt in experimenteel onderzoek naar klanksymboliek. In plaats daarvan maken onderzoekers pseudowoorden (niet-bestaande woorden die dezelfde kenmerken hebben als bestaande woorden, maar geen bekenis) met overdreven klanksymbolische contrasten. Ik probeer drie doelen te bereiken met dit proefschrift: Ik wil er voor zorgen ideofonen een rol gaan spelen in psycholinguistisch onderzoek, ik wil onderzoeken hoe de hersenen klanksymboliek verwerken, en ik wil onderzoeken of er individuele verschillen zijn in hoe gevoelig mensen zijn voor klanksymboliek. Het proefschrift is onderverdeeld in zes experimentele hoofdstukken, zoals hieronder beschreven.

Hoofdstuk 2 geeft een overzicht van het experimentele onderzoek naar klanksymboliek dat tot dusver is gedaan. Het laat zien dat klanksymboliek zich een lange tijd aan de rand van de taalwetenschap en psychologie heeft bevonden, maar dat recente experimenten met pseudowoorden het fenomeen weer populair hebben gemaakt. Dit hoofdstuk roept het onderzoeksveld op om meer interdisciplinair te werken, en om verder te gaan dan het simpelweg observeren van effecten: onderzoekers moeten de effecten proberen te verklaren.

In hoofdstuk 3 heb ik onderzocht of er een verschil is tussen hoe moedertaalsprekers van het Japans klanksymbolische ideofonen en meer arbitraire bijvoeglijk naamwoorden en bijwoorden verwerken. Proefpersonen lazen een hele zin woord voor woord, en moesten daarna beslissen of de zin klopte of niet. Van de zinnen die klopten had de helft een ideofoon en de helft een vergelijkbaar arbitrair woord. Ik heb de hersenactiviteit van de proefpersonen gemeten met EEG (elektro-encefalografie), een methode om de elektrische signalen die de hersenen produceren te detecteren via de hoofdhuid. Ik vond een verschil tussen de twee soorten zinnen. Voor zinnen met ideofonen waren twee aspecten van het EEG signaal anders dan voor zinnen met arbitraire woorden: een sterkere P2 component (een verandering in het signaal rond 250 milliseconden nadat het woord op het scherm verschijnt) en een sterkere late positiviteit (een verandering in het signaal rond 400-800 milliseconden nadat een woord op het scherm verschijnt). Dit suggereert dat de overeenkomst tussen klank en betekenis in ideofonen het makkelijker maakt om klank en betekenis te integreren, vergeleken met gewone woorden.

In hoofdstuk 4 onderzocht ik hoe moedertaalprekers van het Nederlands, die geen kennis hadden van het Japans, Japanse ideofonen en bijvoeglijk naamwoorden zonder klank-betekeniskoppeling leren. Proefpersonen leerden Japanse woorden met de echte Nederlandse vertaling, of met een Nederlands woord dat het tegenovergestelde betekent (bijvoorbeeld kirakira – fonkelend of kirakira – dof) . Ze wisten niet wat de echte vertaling was. De resultaten lieten zien dat proefpersonen de betekenis van ideofonen waarvan ze de echte vertaling hadden gekregen veel beter leerden : proefpersonen scoorden 86.1% correct voor de ideofonen met de echte vertalingen, en maar 71.1% voor de ideofonen waarvan ze een tegenovergestelde vertaling hadden geleerd. Voor bijvoeglijk naamwoorden zonder klanksymboliek was er echter geen verschil: proefpersonen scoorden 79.1% voor de woorden waarvan de ze echte vertaling hadden geleerd en 77% voor de woorden waarvan ze de tegenovergestelde betekenis hadden geleerd. Dit suggereert dat zelfs wanneer je een taal spreekt waarin ideofonen niet voorkomen, je datgene wat ideofonen anders maakt van andere woorden wel kunt herkennen en kunt gebruiken tijdens het leren van een taal.

In hoofdstuk 5 heb ik het experiment van hoofdstuk 4 nogmaals gedaan, maar ditmaal heb ik ook de hersenactiviteit van de proefpersonen gemeten met EEG. De gedragsresultaten van hoofdstuk 5 zijn bijna een exacte kopie van de resultaten in hoofdstuk 4 – proefpersonen scoorden 86.7% voor de ideofonen waarvan ze de echte vertaling hadden geleerd en 71.3% voor de ideofonen waarvan ze de tegenovergestelde betekenis probeerden te leren. Toen de proefpersonen na het experiment gevraagd werd wat ze dachten dat de echte vertaling was van het woord wat ze hadden geleerd, gokten ze de echte vertaling voor 73% van de woorden – veel hoger dan wat je zou verwachten bij toeval. EEG analyses lieten zien dat ideofonen in de conditie waarin de proefpersonen de echte vertaling hadden geleerd, een sterker P3 component (een verandering in het signaal rond 350-500 milliseconden) en een sterkere late positiviteit opwekten. Er was ook een correlatie tussen hoe gevoelig proefpersonen waren voor klanksymboliek en hoe groot het verschil in hersenactiviteit was tussen de twee condities. De resultaten van hoofdstuk 4 en 5 lijken er op te wijzen dat overeenkomsten tussen klank en betekenis het over het algemeen makkelijker maken om woorden te leren, maar dat een conflict tussen klank en betekenis het moeilijker maakt om woorden te leren voor personen die gevoelig zijn voor klanksymboliek.

In hoofdstuk 6 maak ik gebruik van pseudowoorden. Ik heb met een spraaksynthesizer klanksymbolische stimuli gemaakt, om zo de bevindingen van hoofdstukken 4 en 5 in meer detail te onderzoeken . De stimuli maakte ik zo dat soms de klank van het woord overeenkwam met de betekenis (match), soms de klank tegenovergesteld was aan de betekenis (mismatch) en soms daar tussenin (neutraal). Proefpersonen leken gevoelig voor deze gradatie in hoe ze de stimuli beoordeelden, maar bij het leren van de pseudowoorden sprong vooral de match-conditie eruit: het lijkt erop dat het beter leren van klanksymbolische woorden vooral gedreven wordt door de goed passende koppeling van klank en betekenis in de match-conditie. Over het geheel genomen leerdern proefpersonen pseudowoorden beter in de match conditie dan in de neutrale conditie, maar was er geen verschil tussen hoe goed ze woorden leerden in de neutrale en mismatch condities. Maar er waren belangrijke verschillen tussen proefpersonen: 20/60 proefpersonen lieten een gradueel patroon zien (waarbij mismatch het slechtst, neutraal tussenin, en match het beste geleerd werd), voor 16/60 proefpersonen werden juist neutrale woorden het slechtst geleerd, 10/60 proefpersonen presteerden het best in de match conditie en gelijkaardig in de neutrale en mismatch condities, en de overige 14 lieten nog een ander patroon zien. Als we hier het gemiddelde van nemen zien we dat de match-conditie de overhand heeft, maar het illustreert wel dat we meer individuele verschillen in overweging moeten nemen in klanksymboliek experimenten.

In hoofdstuk 7 onderzoek ik individuele verschillen tussen proefpersonen in hoe ze klanksymbolische stimuli beoordelen. Eerdere onderzoeken richtten zich bij mijn weten tot nu toe altijd op slechts één bepaalde taak, bijvoorbeeld met pseudowoorden of juist met ideofonen. Daar is niets mis mee: het helpt om stap voor stap een beter beeld te krijgen van klanksymboliek. Maar wat als klanksymboliek geen homogeen fenomeen is? Wat als proefpersonen verschillende klanksymbolische taken anders doen, en hun prestatie in de ene taak niet gerelateerd is aan hun prestatie in een andere taak? Daarom liet ik proefpersonen in dit hoofdstuk verschillende sets van klanksymbolische stimuli beoordelen – de ideofonen met hun echte en tegenovergestelde betekenissen uit hoofdstuk 4 en 5, de gesynthetiseerde grootte/klank pseudowoorden uit hoofdstuk 6, en een paar woorden waarin er een relatie is tussen de klank van het woord en de vorm van dat waarnaar het woord verwijst. Tijdens al deze experimenten heb ik de hersenactiviteit van de proefpersonen gemeten met EEG. De resultaten van proefpersonen verschilden tussen taken, wat laat zien dat alleen omdat iemand gevoelig is voor één soort klanksymboliek, dat niet betekent dat ze gevoelig zijn voor alle soorten klanksymboliek.

In het discussie-hoofdstuk breng ik de experimentele hoofdstukken samen en discussieer ik de toekomst van onderzoek naar taalsymboliek.

Curriculum Vitae

Gwilym was born in the UK and grew up as a monolingual English speaker, which he spent the rest of his academic career trying to compensate for. He studied Japanese and linguistics at the School of Oriental and African Studies, University of London, receiving his BA in July 2012. He then studied language sciences with specialisation in neuroscience and linguistics (that's not a description, that's the actual title) at University College, London, receiving his MSc in September 2013. He was awarded an International Max Planck Research School for Language Sciences Fellowship by the Max Planck Society, and started work as a PhD student at the Neurobiology of Language Department in September 2013, which culminated in writing the PhD that you've just finished reading. On top of that, he got involved in the Open Access movement as a Max Planck Society Open Access Ambassador, he worked on the *Groot Nationaal Onderzoek* project with Mark Dingemanse, Tessa van Leeuwen, and Linda Drijvers, and he wrote a sort-of-joking-but-actually-serious paper on how academic articles with clickbait-y titles get more attention online.

He now turns numbers into pictures for various organisations as a data visualisation consultant in London.

Publications

Lockwood, G. (2016). Academic clickbait: Articles with positively-framed titles, interesting phrasing, and no wordplay get more attention online. The Winnower, 3: e146723.36330. doi:10.15200/winn.146723.36330.

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(8), 1274-1281. doi:10.1037/xlm0000235.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. Collabra, 2(1): 7. doi: 10.1525/collabra.42.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). Synthesized Size-Sound Sound Symbolism. In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), Proceedings of the 38th Annual Meeting of the Cognitive Science Society (CogSci 2016) (pp. 1823-1828). Austin, TX: Cognitive Science Society.

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioural, developmental, and neuroimaging research into sound-symbolism. Frontiers in Psychology, 6: 1246. doi:10.3389/fpsyg.2015.01246.

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. Frontiers in Psychology, 6: 933. doi:10.3389/fpsyg.2015.00933.

Acknowledgements

First and foremost, I'd like to thank my participants, especially the ones who turned up on time even when it was 9am on a rainy Monday morning. The experiments might have been 90 minutes and some extra beer money to you, but it was three years and an entire livelihood to me. None of this research would have been possible without you.

Many thanks to Mark Dingemanse, my main supervisor throughout this thesis. I was impressed and a little intimidated by your insight, ideas, and intelligence when I first met you, and I still am. Not only that, but I've never met another academic with your enthusiasm, kindness, and consideration towards others. I'm honoured to have been your student, and wherever my future career takes me, I'd be proud to be half the person you are. Thanks also to Peter Hagoort, my promotor, who had the confidence and trust to give me a free rein to go wherever my research took me, and for providing a valuable broader perspective when I was too focused on the details. And thanks to Jyrki Tuomainen, my MSc supervisor at UCL and collaborator on the first experimental chapter of this thesis.

Thanks to my manuscript committee, Steve Levinson, Sotaro Kita, and Suzy Styles. Not just for approving my PhD (although I'm certainly very grateful to you for that), but for poring over it, finding faults, and finding improvements. It was long and frustrating enough to write; I can only imagine how it was to read! Preëmptive thanks to my corona too. As I'm writing this, I don't know who you'll be, but you'll have had to take a few hours out of your schedules to read my whole PhD and then spend an afternoon in a hot aula, possibly while wearing a hat. I really appreciate it.

This PhD might have my name on the front, but its foundations were laid years ago by my inspirational teachers and lecturers. I owe a great deal of thanks to Jim Ferris, Brian Collyer, Sue Motteram, Michi Ashikaga, and especially to Kirsty Rowan and Mandana Seyfeddinipur.

I've enjoyed being part of the NBL group, and I'd like to thank everybody for their input and advice over the years. Special thanks to Ina Grevel, the heart of NBL who makes everything happen. I'd also like to thank Linda, Tessa, and Amanda for their help and collaboration on the synaesthesia projects. Thanks also to IMPRS coordinators Els, Dirkje, and Kevin, the TG, especially Alex, Karin in the library, Jan and Rober, and the admin staff, especially Angela and Marie-Luise.

Thanks to Atsuko Takashima, who kindly voiced all the ideophone and arbitrary word stimuli in my Japanese experiments and became the soundtrack to much of my PhD, and to Arne Driessen for helping with the translations. Grudging thanks to an anonymous reviewer who pulled no punches in their rejection of the manuscript which, after some substantial improvements, eventually turned into chapter 3. You weren't wrong, and I've kept your comments in mind throughout the rest of my PhD.

Thanks to the Language and Genetics group for moving into the new wing. There was no common fridge for milk when I first started, but then you guys moved in, and I could have tea at work again. It's the small things, innit? And as for the large things, many thanks to the EU for making it absurdly easy for a British researcher to be paid through German state and federal government funding to live and work in the Netherlands.

I feel lucky to have met a lot of brilliant people in the Netherlands. I find it quite difficult to write earnest notes of thanks, but *ja, dat bedoel ik*.

Evelien and Elliot, massive thanks for being my paranymphs. I spent most of my time in Nijmegen with you two, and I'm really pleased you'll both be sitting there, ready to jump in with a ridiculous meme in case of emergency. Evelien, sharing an office with somebody for three years could have been a nightmare, but 337 was always gezellig with you. Thanks for being a great office mate and an even better friend. Elliot, there's a time and a place for mucking around, and that was generally in the evening at de Deut (which often felt like my other office). Getting on that wall is still probably the biggest accomplishment of my time in Nijmegen.

Thanks to Lisa for all the parties, Elliott for all the discussions, Suzanne for all the gifs, Franziska and Yves for all the cat cuddles, Richard for all the stats chats, Linda for all the trash TV recommendations, Ilja for all the beer, Sam for all the haircuts (gents, if you're new to MPI, Sam at Daggers will give you the best cut in Gelderland), Amie for all the tea, Monique for all the EEG help, and Johanne, Annika, Will, Ashley, Rick, and everybody else for all the Friday evenings at Cultuur Café. Thanks as well to the MPI and Donders football groups. I hope you remember that one time I scored a rabona rather than the hundreds of times I shanked a shot out for a throw in. Special thanks to Lotte, honorary extra member of 337, for all the kapsalon and all the Dutch translation help (including the thesis samenvatting), and to The Pangolins, Daniel, Dan, and Thijs, for indulging my big musical ideas and mediocre musical skills.

Thanks to my UK friends, Dave, Robin, Rhys, Jonnie, and Lena, who visited me in Nijmegen and saw how life at half the pace is twice as pleasant. And apologies to all my Dutch friends who've invited me round and been offended when I brought beer with me. Not bringing something to somebody else's house was one cultural hurdle I couldn't quite get over.

Thanks to my parents, for many things, but in particular for framing everything as an opportunity, and for encouraging me to try interesting things and see where they lead.

And thanks to Jenny, for everything.

This thesis was written in (on?) Microsoft Word on two laptops and one desktop computer. Writing sessions were soundtracked by 65daysofstatic, Maybeshewill, This Will Destroy You, Caspian, and Tides of Man, and sometimes just ten-hour youtube videos of white noise. It was fuelled by tea and cheap packets of fig rolls.

Max Planck Institute Series in Psycholinguistics

1. The electrophysiology of speaking : investigations on the time course of semantic, syntactic, and phonological processing. *Miranda I. van Turennout*
2. The role of the syllable in speech production : evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed class distinction in spoken-word recognition. *Alette Petra Haveman*
5. The acquisition of phonetic categories in young infants : a self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology : English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited : towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk J. Janssen*
12. Valence and transitivity in Saliba : an Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie H. van der Lugt*
14. Simple and complex verbs in Jaminjung : a study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites : an experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening : the case of phonetic sequences. *Andrea Christine Weber*
17. Moving eyes and naming objects. *Femke Frederike van der Meulen*
18. Analogy in morphology : the selection of linking elements in dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja Helena de Jong*
21. Fixed expressions and the production of idioms. *Simone Annegret Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing : computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing : an electrophysiological approach. *Danielle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra Martine van Alphen*
26. Syllables in speech production : effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Henriëtte Wilhelmina Meeuwissen*

28. Morphology in auditory lexical processing : sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl Jenny Judith Karin Kemps*
29. At te same time.... : the expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension : an electrophysiological approach. *Marijtje-Elizabeth Debora Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Elisabeth Broersma*
35. Retrieving semantic and syntactic word properties : ERP studies on the time course in language comprehension. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren Batya Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency : interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure : distinguishing contrastive from non-constrative variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel Haun*
42. The acquisition of auditory categories. *Martijn Bastiaan Goudbeek*
43. Affix reduction in spoken Dutch : probabilistic effects in production and perception. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition : Electrophysiological evidence. *Valesca Madalla Kooijman*
45. Space and iconicity in German sign language (DGS). *Pamela M. Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages : convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Eva Wagner*
50. Lexical processing of morphologically complex words : an information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo : a Papuan language of the Solomon Islands. *Claudia Ursula Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French : investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*

56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension : evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Payton Robinson*
62. Evidentiality and intersubjectivity in Yurakaré : an interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension : a neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition : insights from the infant brain. *Caroline Mary Magteld Junge*
68. Adjusting to different speakers : extrinsic normalization in vowel perception. *Matthias Johannes Sjerps*
69. Structuring language : contributions to the neurocognition of syntax. *Katrien Rachel Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication : a second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices. The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants' understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sumer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Charles Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntactic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*