Onomatopoeia – Showing-word or Saying-word?

Relevance theory, Lexis, and the Communication of Impressions

1. Introduction

The Oxford English Dictionary defines onomatopoeia as ‘the formation of a word from a sound associated with what is named’. Standard English examples include buzz, meow, crash and splash. Onomatopoeia presents an interesting challenge to the assumption that the link between word form and meaning is completely arbitrary (Saussure, 1916), since the sounds of onomatopoeic words seem to resemble or imitate (at least part of) their interpretations. There is something about the word buzz that resembles the sound a bee makes.

According to Saussure (1916) and other proponents of the ‘arbitrariness of the sign’ doctrine, onomatopoeia is a marginal phenomenon in the study of language, and does not warrant extensive attention in linguistics, the study of language proper. However, as we will show, this is not the case at all. Onomatopoeia is productive, and it is generally acknowledged that many languages have a means of creating words which imitate sounds. Moreover, onomatopoeia raises issues linked to topics which linguists (and pragmaticists, in particular) are very much concerned with, e.g., word coinage, lexicalisation, wordhood, the nature of conceptual meanings, and the communication of
stylistic effects – in particular, those which are impressionistic, indeterminate, and/or non-propositional in nature.

In some languages, and some genres of text or discourse, or in some speech domains, onomatopoeia is quite prevalent. It is has often been noted, for example, that Japanese speakers use onomatopoeia extensively - to the extent that there are dictionaries devoted to cataloguing examples. Similarly, in other, typologically different languages, onomatopoeia is commonly deployed by children’s authors and poets, and is widely encountered in domains of discourse which necessitate reference to the senses, e.g., restaurant reviews, recipe discussions, advertising, and romantic novels. Consider (1) and (2):

(1) He went *galumphing* back.

*Through the Looking Glass* by Lewis Carrol

(2) Noise of crunchy bones goes *crackety-crack* for miles around.

*The BFG* by Roald Dahl

Example (1) contains *galumphing*, which was coined by Lewis Carroll in 1871. Example (2) contains *crackety-crack*. It is very difficult, for both *galumphing* and *crackety-crack*, to describe exactly what these expressions mean. Instead, readers would recover some kind of *impression* that these writers are trying to communicate.

If onomatopoeia is indeed prevalent in the types of contexts mentioned above, then we must consider the nature of these particular contexts in order to shed light on what and how onomatopoeia communicates. We may wish to ask what it is about these particular
communicative situations that leads a communicator to use onomatopoeia in the first place. To achieve a full understanding of onomatopoeia, we must treat it, above all, as a communicative phenomenon.

The best-known approach to the study of onomatopoeia is the sound-symbolism approach, which seeks to find a systematic relationship between sound and meaning. There is also extensive work by Japanese scholars working on onomatopoeia from a grammatical and semantic perspective (Akita 2013a, Kita 1997, 2013, Tsujimura 2001, Toratani 2013). Onomatopoeia has also been investigated in research on synaesthesia (Ward et al 2003, Ward and Simner 2006).

Our aim in this paper is to provide an account of onomatopoeia as a communicative phenomenon. In section 2, we explore the three main existing approaches to onomatopoeia studies in more detail; in section 3, we discuss some aspects of Relevance Theory that make it particularly suitable for the treatment of onomatopoeia, and in section 4, we present our own account.

2. Three approaches to onomatopoeia

2.1 Sound-Symbolism

---

1 Onomatopoeia is also approached in other sub-disciplinary frameworks including the biological anatomy of non-verbal sounds (Assaneo et al 2011) and child language acquisition (Imai & Kita 2015).
Sound-symbolism scholars often assume that there is a systematic relationship between sound and meaning\(^2\). Studies within this framework concern word classes whose sounds seem to have iconic links with sensory experience (Kagitani 2014: 2871), where the link between the phonetic form of a word and its meaning appears to be completely natural and non-arbitrary. It does seem, indeed, that there are situations where some kind of non-arbitrary link between a word’s phonetic form and its interpretation can be observed, whether or not we understand how such links might obtain. The focus in the framework of sound-symbolism studies therefore has been on the nature of the mapping between sound and meaning, and the mechanism(s) underpinning such links.

To illustrate, Ramachandran and Hubbard (2011) repeat a famous experiment by Köhler (1929), showing a jagged shape and a curved shape to college undergraduates, and asking them to decide which shape is ‘bouba’ and which is ‘kiki’. Almost all students judged the jagged shape to be ‘kiki’, and the curved shape to be ‘bouba’. Moreover, Kagitani et al (2014: 2875-2876) report that there is a systematic, statistically significant patterning between the phonetic form and meaning of some onomatopoeic Japanese words for taste and taste texture, e.g., the taste texture ‘thick’ has a relationship to the sounds /n/, /d/, /m/ and /l/, as in neba-neba and doro-doro, while /s/ has a supported link to thinness, as in sara-sara. Furthermore, it is often noted that many words for

\(^2\) The volume of work that takes this approach is too huge to cover in this paper. See, for example, Akita 2009, Hamano 1998, Hinton, Nichols and Ohala 1994, or Ahlner and Zlatev 2010 for detailed discussion.
shininess or light sources in English begin with /gl/, e.g., ‘glow’, ‘glimmer’, or ‘glitter’.

What such studies and observations show is that there are words that clearly seem to have some kind of systematic and non-arbitrary connection to their meanings.

It is unsurprising that we find instances of such words, especially given that the cases studied involve the communication of meanings relating to the senses, and to sensory experience. Humans have rich sensory experiences, and we spend much of our time talking about what we have perceived and experienced through our senses. Many aspects of such experiences, however, are highly idiosyncratic, vague, often ineffable, and impressionistic. As a result, we are unlikely to have lexicalised concepts to express them and communicate them to others. Thus, it is unsurprising that we might try to hit upon some other means to communicate them, using ‘words’ that have a less than fully arbitrary connection to their interpretations. A key question, though, is how and why these links between sound and interpretations obtain in communication.

2.2 The Semantic Status of Onomatopoeia

As we have seen in the previous section, most works on onomatopoeia concern the relationship between sound and meaning, iconicity and onomatopoeia, lexical categories, and word classes of onomatopoeic expressions (see, for example, Akita 2013b, who reviews the development of onomatopoeia research). However, observations have also been made about the relation between onomatopoeia and semantics. For example, Kita (1997: 380) proposes a two-dimensional analysis, arguing
that onomatopoeia/mimetics\(^3\) is independent from other parts of the sentence and works
at the affecto-imagistic level of meaning, where language is directly linked with sensory,
motor, and affective information, while the analytic dimension is ‘the dimension of
decontextualized predication’ (Kita 1997: 380). This distinction is illustrated in (3):\(^4\)

(3) a. * [Taro wa] [isogi-asi de] [haya-aru] o si ta.
   Taro TOP hurried-feet with haste-walk ACC do-PAST
   ‘Taro walked hastily hurriedly’ (lit. Taro did haste-walk with hurried feet)

b. [Taro wa] [sutasuta to] [haya-aru] o si-ta.
   Taro TOP MIM comp haste-walk ACC do-PAST
   ‘Taro walked hurriedly’

According to Kita (1997), (3a) is ungrammatical as the two expressions it contains,
isogi-asi de and haya-aru ki o, both belong to the analytic dimension and are therefore
redundant. In contrast, (3b) is acceptable as it uses an onomatopoeia sutasuta-to,
which belongs to the affecto-imagistic dimension, rather than the analytic dimension.

Tsujimura (2001), in response to Kita’s two dimensional analysis, questions whether we
need a distinction between the affecto-imagistic dimension and the analytic dimension.

\(^3\) Generally, mimetics is seen as involving an imitation of non-sound sensory experience while
onomatopoeia in the strict sense refers to an imitation of sound. We will come back to this point in Section
3 but, for the time being, we use ‘onomatopoeia’ to include both terms.

\(^4\) The abbreviations used in this paper are as follows: ACC = accusative; GEN = genitive; MIM –
mimetic/onomatopoeia; NOM = nominative; QUO = quotative; SUB = subject; TOP = topic.
Tsujimura argues that ‘the meanings of mimetic words are indeed integrated with other parts of a sentence and that they need to undergo linguistic analysis just like other linguistic elements that belong to what Kita calls the analytic dimension’ (Tsujimura 2001: 410). It may be, as Tsujimura suggests, that onomatopoeia (and mimetics) are integrated into purely linguistic structures; however, they might also be like some interjections in this respect – that is, they may be borderline linguistic. We will come back to this point later in 3.3.

2.3 Synaesthesia

As Akita (2013b) notes, synaesthesia has recently started to attract the attention of scholars working on onomatopoeia. According to Simner (2010: 2-3), synaesthesia has been historically understood as involving a sensory or perceptual stimulus triggering unusual sensory experiences or responses. A layperson would consider a typical case of synaesthesia to be something like a mathematician who can count to large numbers using colours and shapes to visualise certain figures. This characterisation of synaesthesia as a ‘crossing over’ of the senses is now considered too broad, and somewhat inaccurate, as much more synaesthesia than previously thought appears to be triggered by graphemes, phonemes and words, so there may be more of a role for linguistic processing in understanding synaesthesia than imagined before (Simner, 2010: 3). Nevertheless, a broad characterisation allows non-specialists in psychology to grasp the phenomenon at hand.
With respect to synaesthesia and onomatopoeia, an important point must be made. Synesthesia is not a communicative phenomenon. It involves generally idiosyncratic links between cognitive domains, which are not within the individual’s control and which not all individuals share. Onomatopoeia, by contrast, is a communicative phenomenon, which can be voluntarily exploited by speakers and appears to work on similar lines across individuals. Nevertheless, looking at onomatopoeia through the lens of synaesthesia can be useful. Examining what happens in a cognitive condition that may involve both sensory and linguistic processing might shed light on how speakers can consciously exploit connections between sensory domains and linguistic processing for communicative effect. Moreover, if it is right that synaesthesia involves multiple senses or perceptual stimuli from different modes, and if it is right that there is a significant involvement of higher level processes bound up with linguistic production and comprehension (Simner, 2010: 3), then we could develop a challenge to the standard view of onomatopoeia as being only concerned with sound.

It is beyond the scope of this paper to discuss this idea in great detail. However, crucially, there is some empirical evidence to suggest that there can be non-arbitrary links between word forms and senses other than sound. In Ward and Simner (2003), a significant and non-random relationship was found between certain phonemes and certain tastes in the mouth: e.g., a phoneme that occurs in the name of a food can trigger a taste corresponding to that food. In Kagitani et al (2014), systematic and statistically significant links were found between particular phonemes and particular
tastes, but also between particular phonemes and particular textures of food in the mouth.

The point is that onomatopoeia need not be dominated by, or confined to, the modality of sound. Many cases of onomatopoeia involve a link between two sounds, but it is likely that connections also hold between the form of a word and other sensory domains. We will discuss such connections involving other senses in section 3.2. This has implications for a key component of how relevance theorists view the ostensive showing behaviours used in ostensive-inferential communication.

3 Towards a Relevance-theoretic Analysis of Onomatopoeia

3.1 Issues with Existing Accounts

These broad approaches to the treatment of onomatopoeia yield a number of insights and empirically supported claims which advance our understanding of onomatopoeia. However, they also raise a number of issues. A chief concern is that virtually every study reviewed for this paper uses examples of established (or fully lexicalised) onomatopoeia, and thus either deliberately or unintentionally restricts the range of data discussed. In fact, many examples are taken from dictionaries, and none of the studies reviewed, with the exception of Hubbard and Ramachandran (2011), aims to investigate what the consequences are from an interpretive perspective of presenting individuals with a novel case of onomatopoeia and asking them what it means. If research into

5 In this study, we will present fairly novel examples, although analyses of the comprehension of nonce onomatopoeias based on empirical evidence would add an interesting dimension to onomatopoeia research.
onomatopoeia does not address why the speaker coined novel or creative forms, how
the hearer interpreted them, and how the expressions became adopted and adapted as
fully-lexicalised words, our understanding of the phenomenon will not be sufficiently
wide-reaching or explanatory. Addressing this explanatory gap is one of the aims of this
paper.

A further issue raised by the approaches taken in previous studies involves the
contribution of onomatopoeia to semantics, understood as the study of linguistically
encoded meanings. Either explicitly or implicitly, some researchers suggest that
particular phonemes (or combinations thereof) encode particular meanings. Thus, many
of these accounts seemingly concentrate on the relationship between sound and
semantics. By contrast, there is no discussion at all of the involvement of pragmatic
processes – in particular the role of pragmatic inferencing – in the interpretation of
onomatopoeia. Furthermore, though several studies (Kita 1997, Toratani 2005, 2013,
Tsujimura 2000) note or imply that there are extra stylistic effects associated with the
interpretation of many onomatopoeias, these studies offer no cognitively grounded
explanation of how these effects obtain. They may present evidence that connections
between sounds and interpretations seem to exist, but the question of how and why
these effects are recovered is overlooked. We propose that this explanatory gap can, in
part, be put down to the lack of a role for pragmatic processes in these accounts.

To fill the explanatory gap, and account for the interpretation of onomatopoeias in terms
of what speakers intend to communicate by them, we need a cognitive account of how
these interpretations are recovered, and how any resulting stylistic effects are derived. We will present our analysis using a cognitive pragmatic framework which is well-equipped to allow us to discuss onomatopoeia in cognitive, communicative, and inferential terms: Relevance Theory (Sperber & Wilson, 1986/1995; Carston 2002; Wilson & Sperber 2012).

3.2 Relevance theory

Relevance theory is centred on two principles of relevance: the Cognitive Principle of Relevance and the Communicative Principle of Relevance. The Cognitive Principle describes how our cognition is organised: human cognition tends to be geared towards the maximisation of relevance. That is, we pay attention to what seems likely to create cognitive effects (improvements to our representation of the world) which are worth our processing effort. You have probably had the experience of buying a new jacket and, the next day, noticing that everyone on campus is wearing the same jacket. It is not the case that everyone on campus bought the same jacket overnight. You notice this because the fact that everyone is wearing the same jacket is now relevant to you: it may strengthen your assumption that this jacket is very ‘in’ for this season, or you may realise that your existing assumption that your dress sense is unique is no longer correct, or you may draw the conclusion that your jacket will be out of fashion soon. In other words, relevance is defined in terms of the balance between cognitive effects and processing effort. Other things being equal, the less effort required to process an input, the more relevant it is. Similarly, other things being equal, the more cognitive effects
derived from processing a stimulus, the more relevant it is. Relevance is a balance between processing effort on the one hand, and cognitive effects on the other.

This cognitive tendency explains how and why communication works: when a speaker makes demands on a hearer’s cognitive effort by producing an ostensive stimulus such as an utterance, the hearer is automatically entitled to assume that whatever the speaker is trying to communicate must be relevant enough to be worth their attention. In technical terms, the hearer, upon recognising this as a communicative act, presumes that the utterance is optimally relevant, and looks for an interpretation compatible with this presumption. This is described in the Communicative Principle of Relevance:

(4) Every act of ostensive communication communicates a presumption of its own optimal relevance (Sperber and Wilson, 1986/1995: 260).

According to this principle, a hearer is entitled to expect a speaker’s ostensive act to be at least relevant enough to be worth processing, and, moreover, to be the most relevant one that the speaker is willing and able to produce at that time.

(5) Presumption of optimal relevance

a. The ostensive stimulus is relevant enough for it to be worth the addressee’s effort to process it.

b. The ostensive stimulus is the most relevant one compatible with the communicator’s abilities and preferences. (Sperber and Wilson, 1986/1995: 270)
The hearer, motivated by the presumption of optimal relevance, then follows a path of least effort in deriving cognitive effects:

(6) The Relevance Theoretic Comprehension Procedure:

a. Follow a path of least effort in deriving cognitive effects: test interpretive hypotheses (reference assignments, disambiguations, implicatures, etc.) in order of accessibility.

b. Stop when your expectations of relevance are satisfied


Relevance Theory (Sperber and Wilson, 1986/1995) was developed to explain how humans communicate particular interpretations through particular ostensive behaviours in everyday face-to-face communication. However, as work in this area has progressed, key aspects of the framework have been applied to other stylistic phenomena, some of which have plenty in common with onomatopoeia, including interjections, expressives, repetitions, appositions, and certain aspects of prosodic behaviour. Given that many onomatopoeias have clear stylistic effects and arguably share characteristics with other stylistic phenomena, it is reasonable to anticipate that existing explanations and theoretical notions used to explain these other phenomena can be co-opted to shed light on onomatopoeia.
Let us start with the relevance-theoretic notion of ‘representation by resemblance’. Sperber and Wilson (1986/1995) show how communication is not purely a matter of describing the world, and how resemblances of all types – visual, auditory, phonetic, linguistic, semantic, topological – can be exploited in communication. This aspect of communication is generally overlooked by other pragmatic frameworks. As Sperber and Wilson (1986/1995: 227) explain, ‘in appropriate conditions, any natural or artificial phenomenon […] can be used as a representation of some other phenomenon that it resembles’. For example, when offered a glass of wine, one could raise both hands to chest height and move them as if driving, to produce a communicative act plausibly interpreted as a refusal. Two objects or actions resemble each other to the extent that they share properties. The hearer, following the relevance-theoretic comprehension procedure, must determine which properties are to be identified as shared. The most salient ones yielding enough implications about the stimulus (for example, drinking and driving) will be those that make the communicative act worth the hearer’s attention.

Relevance Theory also provides an explanatory framework that accounts for descriptive ineffability, or expressive effects (e.g. Blakemore 2008, 2011, 2015, Wharton 2009). Not all communicative acts have the goal of achieving a single strong cognitive effect or a small, determinate array of effects. As Sperber and Wilson (1986/1995) explain, communication is a matter of degree: some utterances are intended to trigger the recovery of a single, strongly evidenced proposition, while others communicate a broader array of weaker assumptions, and, in some cases, what is communicated amounts to no more than an impression, which is hard to render in propositional terms.

For example, Blakemore (2008), in analysing the use of apposition in a variety of texts, argues that some cases of apposition give rise to apparently non-propositional effects, effects so intangible and ineffable that they cannot be paraphrased without loss of meaning. As a result, word-by-word translation or attempts at paraphrasing such phenomena inevitably destroy some of these effects. According to Blakemore (2011), phenomena such as expletives, epithets, diminutives, and interjections typically communicate expressive effects of this type.

Apart from these specific notions, Relevance Theory offers a number of fundamental assumptions that make it extremely well-suited to developing an account of onomatopoeia. First, there are cases of onomatopoeia which are fully lexicalised. These lexicalised expressions must originate somewhere, and this suggests that there must be plenty of novel, creative and even ‘one off’ cases of onomatopoeia which are neither coded nor conventionalised and may therefore have something in common with non-verbal communicative behaviours such as affective vocalisations, gestures, and facial expressions. Onomatopoeia should therefore be handled within a framework that can account for the role of both verbal and non-verbal behaviours in communication. A key assumption in Relevance Theory is that both verbal and non-verbal behaviours can be equally communicative, and neither mode has a privileged status within the framework.\(^6\)

This is possible because pragmatic inference is seen as playing a crucial role in the interpretation of all communicative behaviours, both verbal and non-verbal. From the

---

interpretation of the most commonplace of lexical items such as *red* or *drink* to the
interpretation of spontaneous, non-coded, non-verbal behaviours, audiences must infer
what these communicative behaviours are intended to convey, how they are to be
optimally processed, and what type of effects it is relevant to derive in particular
contexts. This framework enables us to bring inference into the interpretation of
onomatopoeia and account for how its effects are derived in a wide variety of cases, but
in particular, in the novel and creative cases which, crucially, must be interpreted
entirely inferentially because the audience is not familiar with any code or convention
governing their use.

Relevance Theory is not the first pragmatic approach to acknowledge the role of both
verbal and non-verbal behaviours in communication. For instance, Grice (1957) was
much concerned with establishing a distinction between *showing* (which is typically non-
verbal) and *saying* (which is typically verbal) in developing his theory of speaker's
meaning. Thus, I may point out of the window to *show* you that it is raining, or I may use
the words ‘*Il pleut*’ to *say* that it is raining. However, Relevance Theory (see, for
2003, 2009) has provided good reason to think that there is a continuum rather than a
distinction between *saying* and *showing*, and that a theory of communication should
encompass both elements. This makes it easier to account for cases involving a mixture
of both *showing* and *saying* (e.g., saying ‘I’m angry’ in an angry tone of voice). We will
argue that lexicalised onomatopoeia provides many cases of this type.
Grice (1957) introduced another important theoretical distinction, between 'natural' and 'non-natural' meaning. Natural meaning is carried by so-called 'natural signs' (e.g., smoke is a natural sign of fire), and non-natural meaning involves the overt expression of a communicator's intentions. Wharton (2003a, b) argues that natural signs can also be exploited in overt intentional communication. As he points out, utterances are often accompanied by non-verbal 'cues' such as facial expressions, tones of voice, gestures, bodily movements, etc. Often, these non-verbal cues are not intentionally provided, and the communicator need have no control over them. Suppose you have the flu. Hearing you cough, a hearer might conclude that you are unwell and need to go home. This is a case of accidental information transmission, which provides the hearer with information whether or not you want him to know you are unwell. Unless you make clear that you intend to communicate that you are unwell, it would not be a case of overt intentional communication (or ostensive communication, in Relevance Theory's terms). However, as Wharton (2003a, b) argues, there are cases in which the speaker intentionally exploits such natural behaviour to provide overt evidence for an intended meaning. For example, if you cough suggestively in front of your friend, then you have just produced a natural behaviour to give your friend overt evidence that you are unwell and need to go home soon.

According to Relevance theory, when we communicate, there are two layers of information that the hearer must retrieve. As Sperber and Wilson (1986/1995: 51-54) explain, the first layer is the information that the speaker intends to communicate, and the second layer is the information that the first layer of information is being pointed out.
intentionally. Recall the earlier example of coughing. By coughing suggestively in front of your friend, not only do you make manifest to the hearer the first layer of information, i.e. that you are unwell, but you also make manifest to the hearer the second layer of information, i.e. that you intend to inform him that you are unwell. That is, ostensive-inferential communication involves making manifest to the hearer that you intend to make the first layer of information manifest. As Wharton (2001) argues, exploitation of non-verbal cues in ostensive-inferential communication is a case of overt intentional showing rather than saying.

The difference between showing and saying is generally analysed in terms of the ‘directness’ of the evidence provided for the first layer of information, where evidence derived via linguistic decoding and inference is relatively less direct, requiring more of an ‘inferential leap’ from evidence to intended interpretation. In cases of showing, the ostensive behaviour of the speaker provides direct evidence for the first layer of information. If you receive a gift and wish to communicate disappointment with it, you can allow people to see your frown, a natural sign of disappointment. Your disappointment is directly inferrable from your frown. Or you can provide less direct evidence for the first layer of information by saying ‘I am disappointed’, from which the hearer arrives at the first layer of information by a combination of linguistic decoding and inference.

Cases of showing often arise when a communicator wishes to communicate something that is difficult to put into conceptual terms using language. It may be that what the
communicator wants to convey is too vague and ineffable, such as a complex feeling or sensory experience. Or, it may be that providing direct evidence is more convincing (e.g., pointing out black clouds to communicate that you expect it to rain). It seems reasonable to say that *showing* often involves the sharing of experiences, or draws on experiential elements of the context.

Addressing the role of interjections in communication, Wharton (2000, 2009) relies on a distinction between *saying*, where the speaker provides indirect and coded evidence (e.g., utterances), and *showing*, where the communicator provides direct (and often natural) evidence for what he wants to communicate. Wharton (2000, 2009) demonstrates how the continuum between *showing* and *saying* works using interjections and other expressions that straddle the border between natural and non-natural meaning. However, it is worth noting here that most examples of *showing* in the relevance-theoretic literature involve cases of *showing* by bodily expression, or by presenting visual evidence to the hearer. Wharton (2009) uses gaze and ostensive sighing. Similarly, Sperber and Wilson (1986/1995) use a range of examples that provide direct visual evidence, e.g., showing a bottle of aspirin to communicate that someone is unwell. Metalinguistic resemblance, that is, exploitation of resemblances in linguistic form, has also been discussed extensively (e.g., Sperber and Wilson 1986/1995, Noh (2000), Wilson and Sperber 2012). The case of onomatopoeia adds an interesting extra dimension to the discussion of *showing*, as the evidence it provides is not properly linguistic, but onomatopoeias are not bodily or visual either. Instead,
onomatopoeia exploits resemblances between phonetic forms and events in the world.

In a sense, *showing* by onomatopoeia is a case of cross-modal *showing*.

Within pragmatic stylistics and within Relevance Theory, there has also been very little work on onomatopoeia. To our knowledge, the only treatment of it within Relevance Theory is a brief but very promising discussion in Wharton’s (2009) analysis of interjections. Wharton (2009: 99) considers that onomatopoeia, and related cases of iconic language use in general, involve an interaction between coding and inference processes at the lexical level, where ‘words’ which are arguably linguistic in nature seem also to exhibit a degree of *showing*, i.e., to provide more direct evidence for what the speaker intends to communicate. We use these insights to inform our analysis of a range of cases of onomatopoeia, running from those which are novel, natural, and probably characterisable solely in terms of *showing*, through a middle range of terms which combine *showing* and *saying*, through to cases which are so fully lexicalised that there is essentially no *showing* to be identified, and there is only a trace connection to the sensory domain that triggered the original onomatopoeia.

In the next section, following Wharton’s (2003, 2009) discussion, we explore similarities and differences between onomatopoeia and interjections.
3.3 Onomatopoeia and Interjections

Intuitively, we might expect to find similarities between onomatopoeia and interjections. After all, if it is right that some cases of onomatopoeia behave as words which are not (fully) integrated into a language, then it is reasonable to expect to be able to make some comparisons between them and interjections, another case of expressions which are generally considered to sit at the edge of the linguistic systems of natural languages, (e.g., Goffman 1981, Trask 1993, cited in Wharton 2003, 2009). In this section, we follow Wharton (2003, 2009) in treating interjections as expressions which can show what the speaker wants to communicate, or simultaneously show and say what the speaker wants to communicate. We establish that, syntactically, there are no grounds for treating onomatopoeia and interjections as facets of the same phenomenon. Thus any similarities between onomatopoeia and interjections must lie, if anywhere, in what they communicate, and how. That is, they must arise from the way a speaker uses both types of expression to provide evidence for the information he or she wishes to communicate, and the nature of the information communicated, i.e., vague, indeterminate, impressionistic, or non-propositional effects.

According to Wharton (2003: 175), interjections are generally considered to be at the edge of language; historically, linguists have regarded them as a paralinguistic, or even

---

7 To be clear on the sort of expressions we have in mind here, the following are standardly considered to be interjections of one type or another: wow, ouch, argh, yuk, hell, damn, and shit.

8 Some scholars (see, for example, Meinard 2015) suggest that some cases of onomatopoeia might be analysable as a special type of interjection.
completely non-linguistic, phenomenon. They have even been described as mere
decoration (Sapir, 1970, cited in Wharton, 2003: 175). From a syntactic perspective, it is
often claimed that interjections proper are not syntactically integrated into their host
utterances (Trask, 1993), and that they always constitute an intonation unit of their own
(Ameka, 1992), e.g.:

(7) Shit! | I've lost my prescription.

(8) That | – ow! – | really hurts.

However, there are no syntactic grounds for drawing a comparison between
onomatopoeia and interjections. Recall Tsujimura (2001), who argues that
onomatopoeia is fully integrated into linguistic structures. Consider (9) and (10):

(9) | Meow! | The little cat said. | Buzz! | The little bee went. | Splash! | Went the sea
lion.

(10) | Buzz is the noise a BEE makes |, and cats say meOW. |

Standard onomatopoeic expressions such as meow, buzz, and splash \(^9\) constitute their
own intonation groups in (9), but are syntactically (and prosodically) integrated in (10).
Even if interjections are considered to be syntactically isolated, many onomatopoeias
clearly are integrated into the utterances that host them. There are thus no syntactic

\(^9\) As we will see later, onomatopoeic words seem to exhibit different degree of lexicalisation. Classic examples of onomatopoeia such as splash are considered more lexicalised than stylised imitations such as meow and buzz.
grounds for drawing parallels between onomatopoeia and interjections. Instead, we have syntactic grounds for saying that onomatopoeias do not necessarily, at least in a syntactic sense, exist at the edge of language proper. In addition, most onomatopoeias inflect morphologically, in a way that true interjections do not.

What, then, do interjections communicate, and how? From the very broad perspective of the type of interpretations they give rise to, there is a consensus that interjections are used to communicate emotions and attitudes (Wierzbicka, 1991; Wharton, 2003; Meinard, 2015: 151). What is difficult to explain is how such expressions communicate something as idiosyncratic and ineffable as an emotion or an attitude.

On what is generally described as the ‘conceptualist’ approach\(^\text{10}\), interjections are considered to have a semantics (and, thus, from a semantic perspective, to be part of language) because they have conceptual content (Wilkins, 1992: 119). On this approach, they are analysed as communicating complex conceptual structures, and as doing so by encoding them. These conceptual structures are seen as composed of irreducible semantic primitives similar to those in the Natural Semantic Metalanguage (NSM) (Goddard, 2011), e.g., YOU, THIS, GOOD, THINK, DO, BE, DIE, PLACE. Wierzbicka (1992: 164) proposes the following conceptual structure for wow:

---

\(^{10}\) There is a great deal of work on the conceptualist approach to interjections. Because of space limitations, we cannot address all of it, particularly with respect to the distinction between primary and secondary interjections (Wilkins 1992, Ameka 1992). The aim here is merely to provide enough discussion to suggest that many interjections do not linguistically encode anything conceptual, and to argue (as Wharton, 2003 does) that what some interjections communicate is the type of thing that is far too nebulous to be linguistically encodable in conceptual terms.
I now know something
I wouldn’t have thought I would know it
I think: it is very good
(I wouldn’t have thought it could be like that)
I feel something because of that.

A key objection here is that decompositional attempts to define interjections invariably fail, as Wharton (2003, 2009) points out. Whilst it is possible to define some words in conceptual terms, the attempt to supply sets of necessary and sufficient conditions for most or all words generally fails (Fodor, Fodor and Garrett 1975, cited in Wharton, 2003: 178). Wharton (2003) argues that expressions such as wow can be used to communicate a wide range of emotions and attitudes, some of which are negatively oriented. To accommodate such negative emotions, we would need to make the conceptual structure even more complex. Moreover, as he points out, not every element of such a structure, however complex, is always needed – one can felicitously utter wow without thinking one would never have expected to know something, for example. Thus, it is difficult to find a set of necessary and sufficient conditions that would allow us to define interjections in conceptual terms. Attempts to paraphrase as a set of propositions what an interjection communicates always lose something ‘in translation’. Compare (12) and (13):

(12) Yuk! I hate worms.
I am disgusted. I hate worms.

Yuk and I am disgusted do not yield equivalent interpretations, and the latter is paraphrasable while the former is not. Moreover, while I am disgusted can be true or false, Yuk! has no truth value at all. It then seems reasonable to conclude, as Wharton does, that interjections do not encode concepts. However, and this will be important in our analysis of onomatopoeia, interjections do communicate something. For example, wow can communicate a potentially indefinite range of emotions and attitudes depending on the context and the intentions behind its use. Since interjections cannot be seen as encoding concepts, a different analysis of how they communicate is required.

Drawing on work by David Kaplan (1997), Wharton (2003, 2009) points out that interjections are typically used to reveal something about, or express, the speaker’s emotional states, and that they fall on the expressive rather than the descriptive side of language use. He concludes that while interjections cannot be seen as encoding concepts, they might be analysed in procedural terms. According to Relevance Theory, two types of information can be encoded: conceptual, and procedural (see, for example, Blakemore 1987, 2002, Wilson and Sperber 1993). Instead of (or as well as) encoding concepts, an expression may encode procedures – instructions for how an utterance is to be interpreted, or for activating particular kinds of representations or effects – just as if a button is being pushed (Blakemore, personal communication). Having argued that interjections do not encode concepts, Wharton (ibid.) argues that they encode
procedures, which may activate representations of a wide range of emotional or attitudinal states.

As mentioned earlier, Wharton (2003, 2009) analyses interjections in terms of the showing – saying continuum. What he suggests is that interjections can be treated as expressions that show and say simultaneously. Being partly coded, and partly related – via the use of affective intonation, for instance – to something like natural cries, we might want to analyse them as contributing to an interpretation in two ways: by simultaneously showing and saying, as in (14):

(14) Ann: Don’t forget, you said you’d proofread Kelly’s essay tomorrow.
   David: Damn!

In (12), we might say that David, uttering ‘damn!’, both shows and says (because of the coded element in the interjection) that he is annoyed at having to correct the essay. The coded element in the interjection (or the procedure encoded by the linguistic form) activates a particular kind of representation (i.e., of states associated with annoyance) while the showing element in the interjection provides the direct evidence for the first layer of information to be picked up.

In arguing that interjections can show and say at the same time, Wharton (2003: 206) mentions onomatopoeia as supporting evidence. He says that clink, splash, and sizzle are examples of words which not only encode concepts but also seem to show
something about the objects they denote, since the link between the sounds of these
words and their meanings is not completely arbitrary. Thus, *splash* might encode a
concept SPLASH, but there is still an element of *showing* in that the sound of the word
bears more than a passing resemblance to the sound of the phenomenon it denotes.
This is perhaps because onomatopoeic expressions communicate elements of sensory
experience, feelings, or impressions, which are vaguer and harder to pin down in
conceptual terms than the meanings of purely conceptual expressions like *dog* or *green*.
It would thus be reasonable to hypothesise that the ‘meanings’ of onomatopoeic
expressions range from fully established concepts to apparently ‘non-propositional’
effects, that they do have a conceptual semantics in some cases, but that other aspects
of their meaning fall more on the *showing* than the *saying* side.

Following Wharton (2009), we adopt the idea that there is a continuum of cases
between *showing* and *saying*. Accepting that onomatopoeia can *show* and *say* at the
same time allows us to consider a continuum of cases from pure *showing*, through
mixed *showing* and *saying*, to just *saying*, enabling us to account for the disparate
range of completely novel to fully lexicalised cases of onomatopoeic expressions. And
indeed, some researchers assign degrees of ‘lexicality’ (or lexicalisation) to
onomatopoeias, which can be determined based on their possibilities of inflection and
occurrence in quotation. For example, Kadooka (2005) illustrates degrees of lexicality
ranging from the least lexicalised expressions, such as the English *hjckrrh* (an
exclamation used by Lewis Carroll in *Alice in Wonderland*) and Japanese *bakyuuun*
(sound of a gun), through middle cases, such as *meow* and *pop* in English and *karari*
(onomatopoeia for dryness) and sowa-sowa-suru (nervous) in Japanese, to the most lexicalised cases such as English chatter and Japanese odoroku (surprised). Flyxe (2002) also lists expressions such as kippari (clearly) and odoroku as examples of highly lexicalised onomatopoeia, and expressions such as zabun (splash), or gān (an expression often used when someone is shocked) as among least lexicalised onomatopoeia. It seems reasonable to assume that the least lexicalised cases, such as hjckrrh and gān, exist at the showing end of the spectrum while chatter or kippari are located at the saying end.

Onomatopoeias have iconic features not because there is a form/sound-meaning relationship of the type described earlier in this paper, but because what is being communicated is information about what the speaker perceived. As Wharton (2009: 100) argues, ‘in onomatopoeic expressions generally, the link between sound and meaning is not as loose as in most other words, since some element of the natural connection remains’. If a speaker wants to communicate what she perceived, she should use what she thinks is a faithful enough representation of it, that is, one with some natural cognitive resemblance to the original perception. According to Wharton (2009: 101), this suggests that the showing-saying continuum applies not only at a non-lexical (non-verbal) level but also at a lexical level, contributing to lexical-pragmatic processes.

In the next section, we present our relevance-theoretic account of onomatopoeia, using examples ranging from relatively creative onomatopoeias to more established ones. In
particular, we explain what onomatopoeia communicates, how it communicates, and why the speaker chooses to use onomatopoeia rather than fully lexicalised, non-onomatopoeic, conceptual terms.

4. A Relevance-Theoretic Approach to Onomatopoeia

4.1 The Showing-Saying Continuum and Onomatopoeia

So far, we have seen how onomatopoeia is analysed in various fields and have demonstrated how little research there has been on the role of onomatopoeia in communication. Against this backdrop, we treat onomatopoeia as a communicative phenomenon and present an alternative explanation of the role it plays in communication. Drawing on Wharton’s (2009) account of interjections, we argue that onomatopoeia falls on the showing-saying continuum. At the showing end of the spectrum, the communicator uses novel (creative) onomatopoeias, such as hjckrrh (Kadooka 2005), that are not established as words per se, to share his sensory experience similarly to the way non-verbal behaviours can be used to ‘show’ what a person intends to communicate. At the other end of the spectrum, where the communicator uses more established onomatopoeic words, such as chatter, there will be less of a showing element, and more of a saying element. What is interesting, though, is the middle range of cases, such as pop or sizzle in English, and sowa-sowa or karari in Japanese, where onomatopoeic words seem to have characteristics of both showing and saying. This is where onomatopoeia exhibits an interaction between coding and inference at the lexical level.
Why does a communicator use onomatopoeia and what does he want to ‘show’ rather than communicating via purely conceptual encoding? These questions are related to a very fundamental question: what does onomatopoeia communicate? We have argued that onomatopoeia can *show* and *say* at the same time, and can communicate a very vague, apparently non-propositional, interpretation, which, in relevance theory, has been called an *impression*. An impression is a sub-type of cognitive experience: one that involves a diffuse range of evidence, often sensory or emotional, pointing towards a certain conclusion, or a certain answer to a question (e.g. from the way someone talks, I get the impression that he’s nervous; from looking out of the window, I get the impression that it’s going to rain; from the word ‘sizzle’, I get an impression of what sausages sound like when cooking).  

When reporting impressions from his sensory experience, the speaker ‘imitates’ the experience using the phonemes available to him. In order to create the most faithful representation possible of the original experience, he would use sounds that most closely resemble the impressions of that experience that he wishes to communicate. Onomatopoeia is not purely ‘natural’ in the same way that some interjections can be as a physiological response to stimuli. Onomatopoeia is a somewhat stylised and iconic representation of sensory experience via sound, whether the original experience was aural, visual, or of any other sensory type. In other words, onomatopoeia is what the

---

speaker considers a faithful enough representation of the sensory experience that he wishes to share with the audience.

The alleged systematic link between onomatopoeia and sound comes from this. The sound of each onomatopoeia triggers access to certain qualities often associated with such sounds (voiceless sound = clarity, for example). However, this does not mean that each sound 'encodes' the relevant quality. The triggering is context-dependent, and the quality is accessed only in some specific contexts, and not others.

Let's take a relatively established onomatopoeia, *sizzle* in English and *sowa-sowa* in Japanese. Even when these expressions and their meaning are fully established, there is still an element of *showing* involved, leading to the communication of a sensory impression as part of the first layer of information the audience is intended to pick up.

(15) Sausages are *sizzling* in the pan.

(16) [Taro wa] [sowa-sowa] shiteiru.

TaroTOP MIM do-PRESENT

‘Taro is nervous.’

While the concept encoded by *sizzle* denotes making a certain type of hissing noise, the word still evokes something more specific by use of the /s/ sound, giving an impression of the kind of sound that the sausages make. Similarly, the concept encoded by *sowa-sowa* denotes a certain type of nervousness, although the word still communicates an *impression* of the kind of nervousness exhibited by Taro, presumably by (the repetition

31
of) its phonemes. Thus, using onomatopoeia enables the speaker to communicate not just encoded concepts, but an additional sensory impression which is extremely difficult to put into words. While well-established, lexicalised onomatopoeias may have a stronger saying element\textsuperscript{12}, their phonetic link to the senses enables the communicator to include some expressive effects.

4.2 Onomatopoeia and Sensory Experience

Let us move on to a case of creative onomatopoeia. There are a number of onomatopoeias in Japanese commonly used for wind: *pyu-pyu*, *byu-byu*, *hyu-hyu*, *hyuuu*, *sa-*-, *soyo-syo*, and *suu-suu*. However, in (17), the sound *do*, which is not commonly used to describe wind, is used in many variations:

(17) [Describing the strong wind that is storming through the village]

\begin{verbatim}
Doddodo dodō do, dodoōdo, dodō aoikurumimo fukitobase
MIM blue chestnuts also blow-away
Suppai Karin mo fukitobase Doddodo dodō do, dodoōdo, dodō sour quince also blow-away MIM
"Doddodo dodō do, dodoōdo, dodō; Blow away the green chestnuts too;
Blow away the sour quince too; Doddodo dodō do, dodoōdo, dodō...."
\end{verbatim}

\textsuperscript{12} For this reason, it may be plausible to argue that highly lexicalised onomatopoeia contributes to ad-hoc concept construction.
Not only does this example contain creative onomatopoeia, it also has a rhythm that is different to the repetition of two syllables often seen with Japanese onomatopoeia\(^\text{13}\).

That is, the sound of wind in (17) is expressed in a multimodal manner via the combination of sound and rhythm. This is particularly effective for communicating the sensory experience evoked in this scene. The voiced plosive /d/ sound, which is often associated with heaviness, evokes the sense of a heavy and strong wind, while the rhythm communicates the somewhat musical, continuous movement of the air. The multimodal nature of onomatopoeia provides evidence that the link between a particular sound and ‘meaning’ is determined by the way the communicator perceives the particular sensory experience in the context of communication, and not because there is a non-arbitrary link between sound and meaning\(^\text{14}\).

Earlier, we mentioned that onomatopoeia communicates sensory experience via sound. This idea may enable us to distinguish between onomatopoeia in the strictest sense, and mimetics. Typically, as noted above, onomatopoeia is defined as an imitation of sound, while mimetics is defined as mimicry of non-sound. Thus, onomatopoeia, in the

\(^{13}\) The repetition of two syllables is also common in English diminutives e.g., doggy woggy, potty wotty, daddy waddy. We wish to thank an anonymous reviewer for these examples.

\(^{14}\) It is also interesting to see how these onomatopoeias are left untranslated in (17). The translator, perhaps, felt translation of this onomatopoeia into another linguistic form would not achieve similar effects.
strictest sense, is a case of *showing* within the same sensory domain, while mimetics is a case of cross-modal *showing*, where the speaker uses sound as a medium to express a sensory experience from a different sensory domain. Note that it can be very difficult to determine which sensory organs a particular onomatopoeia is linked to. The cross-modal nature of *showing* by the use of onomatopoeia (in a broad sense, including mimetics) could account for the complex layers of expression in onomatopoeia.

The use of novel and cross-modal onomatopoeias is nothing special. Example (18) is taken from a children’s picture book about a little girl who is out in the field picking flowers. While she picks flowers, a number of animals fall from the sky:

(18a)

![Figure 1: Scene containing onomatopoeic expression for a fallen crocodile: dozuzun](image)
Figure 2: The use of onomatopoeic expression for a fallen elephant: *dokashiin*

(18c)

Figure 3: The use of onomatopoeic expressions for a fallen zebra, lion and panda:

*guwashi, bako, dongorogorogorogoron*

(18d)
Figure 4: The use of onomatopoeic expression for a fallen mother: *kuru, suta*

(All examples from Izumi Motoshita and Kiyotaka Ishii, *Futtekimashita* [falling down] 2007)

Examples (18a) to (18c) denote scenes where heavy objects (i.e., animals) fall from the sky.\footnote{We are most grateful for an anonymous reviewer’s comment on these examples.} The standard onomatopoeic expression for fallen objects is *dosu(n) / dosa / doka(n) / doshi(n)*. However, while examples (18a) to (18b) all include /d/, /b/ or /g/ sounds which are often associated with the sound of heavy objects falling, none of them uses standard onomatopoeia for fallen objects. Instead, a variety of onomatopoeic expressions that are made up from parts of established onomatopoeia are used. (18) seems to be the combination of *doka* and *doshin*, while *dozuzun* in (18a), possibly related to an established onomatopoeia for falling objects, *dosun*, seems to be associated with the way the action was continuous (or lasted longer than a common instantaneous fall). Moreover, (18c) includes *gorogoro*, onomatopoeia for the manner of objects rolling, as well as a variation of *gashi* and *bako*, which are often used for a
collision with heavy objects rather than a fall. The departure from highly lexicalised onomatopoeia and the use of creative combinations of these sounds associated with qualities in different modalities enables the communicator to convey different impressions of different multi-sensory experiences. Example (18d), in contrast, does not contain /d/. Instead, it contains the voiceless sounds /s/ and /k/, presumably to distinguish the different instances of landing. It also has much a shorter duration, indicating the lighter and more agile manner in which the event took place, where the mother successfully landed like a gymnast. Examples such as these pose problems for the sound-symbolsm approach, as it is not clear if deviations from standard onomatopoeic expressions and creations of new ad hoc ones would require new categories. After all, most works taking the sound-symbolsm approach concentrate on the iconicity between ‘sense / perception’ and sounds, while seemingly overlooking communication. To some extent, as we have repeatedly said, it is not surprising that the link between perception and sound is not arbitrary – humans attempt to ‘recreate’ what they perceive, using the tools available to them, which happen to be language and, hence, onomatopoeia in these cases. This results in some resemblance and, hence, the link between sound and meaning does not appear ‘arbitrary’. However, this shouldn’t mean that we can pinpoint the ‘meaning’ of sound; the same sound occurs in a variety of contexts and the interpretation of such onomatopoeia is context-dependent. Nor does it necessarily mean that this alleged link between a sound and its interpretation in a specific context is linked to the arbitrariness of language. In fact, it does not matter to the current study if the link between verbal sound and linguistic meaning is generally
arbitrary or not; our aim is to explain how speakers use onomatopoeia, and how hearers recover their intended interpretations.

As we have seen, it is claimed in Relevance Theory that communication is a matter of degree, and that not all ostensive acts convey a single strongly-evidenced proposition. In some cases, the intended interpretation is an array of weakly evidenced assumptions, or even what appears to be a wide array of intangible, non-propositional effects. This is what (at least, less-established) onomatopoeias communicate.

Onomatopoeia, as a tool for showing the communicator's sensory experiences, enables the speaker to communicate the impression she experiences in a particular situation. The (non) arbitrariness of the sound-meaning relationship matters little when it comes to what is actually communicated. Whatever the link /b/ or /u/ has with our cognition (or what we perceive in different domains of cognition), the fact is that the speaker uses them as a tool to communicate an intended interpretation that will suit the particular context. In other words, these expressions are just a pointer towards the non-propositional perceptions of the speaker.

4.3 Context Dependency and Interpretation of Onomatopoeia

In (18), we saw how different expressions are used to represent similar sensory experiences. In contrast, as example (19) shows, the same sound don, for example, could be used to represent totally different manners of performing different actions:

(19a) kare wa dondon to doa o tataita.
‘He banged the door.’

(19b) Iroirona koto ni dondon chosen-shite hoshii.

‘I’d like you to challenge various new things.’

While dondon in (19a) is a stylised imitation of the sound of banging a door, in (19b), it is used to denote a more abstract concept: how a person deals with life. Examples (20) and (21) demonstrate similar context-dependency of the interpretation of onomatopoeia.

Tsujimura (2001: 45) lists how burabura seems to convey a range of meanings as seen in (20), and Mikami (2004: 3) reports similar elusiveness for gorogoro, as shown in (21):

(20a) Doa no totte ga burabura-suru.

‘The doorknob is loose.’

(20b) Ashi o burabura-si-naide suwarinasai.

‘Sit without swaying your legs.’

(20c) kooen o burabura-sita

‘I strolled in a relaxed way in the park.’
(20d) Otto ga uti de burabura-site iru.

husband SUB home LOC MIM-is doing

‘My husband is wasting time at home (without doing anything important).’

(21a) [The sound of thunder, or a thunder-like sound]

Enrai ga gorogoro-to narinagara dandan chikaduite kuru youda.

far-thunder SUB MIM-QUO roaring dandan approach-come looks-like.

‘It looks like thunder is gradually approaching.’

(21b) [The manner in which heavy objects or bodies roll in sequence]

danborubako o katamukeru to, migotona jagaimo ga gorogoro korogarideta.

cardboard box ACC tilt case impressive potatoes SUB MIM roll-out PAST

‘When we tipped the cardboard box, impressive-looking potatoes came rolling out.’

(21c) [To spend time without working or doing anything particular]

shisshoku shite inaka no oyamoto ni kaeri, ichinen hodo gorogoro-to kurashiteita.

lost-job did hometown GEN parents to return, a year about MIM-QUO lived.

‘(I) lost my job, went back to my parents, and lived doing nothing for about a year.’
(21d) [The way an item does not have a special quality]

ano teido no bijin nara, Tokyo ja gorogoro iru yo.

That degree GEN beauty if, Tokyo in MIM exist SF.

‘Beautiful ladies of that level will be plenty in Tokyo.’

(21e) [The feeling of discomfort when a foreign item enters]

gomi ga haitte me ga gorogoro suru.

rubbish SUB enter eye SUB MIM do.

‘Something got into my eyes and it hurts.’

These examples show how the same sound can communicate similar ‘meanings’ or different ones, raising questions for the sound-symbolism approach.

So far, we have seen examples where the differences in meaning are relatively obvious. However, differences in what onomatopoeia communicates can be very subtle. Examples (22) to (24) show the use of the Japanese onomatopoeia for silence:

Figure 5: The use of onomatopoeia for silence in picture book
Mori no naka wa shin to shite kooru yona samusa desu
Forest GEN inside TOP MIM QUO do freeze as-if coldness COP

‘It is very quiet and freezing cold in the forest.’

Figure 6: The use of onomatopoeia for silence in manga
Figure 7: The use of onomatopoeia for silence in action manga

(ONE and Yusuke Murata, One-Punch Man, Volume 4 2012: 24-25)

Example (22) is taken from a children’s picture book. In this scene, animals visit the forest with Christmas gifts for the trees. Here, shin is used to communicate the impression of a quiet morning in a winter forest. In (23), a scene taken from manga, a long-vowelled version shi-n is used with ‘silence’ as the English annotation. Here, the characters, who had been gossiping about the protagonist, fall silent when they realise he is within earshot. Shi-n is also used in (24), this time to describe the disappearance of the enemy. What is interesting is not the fact that (variations of) shin are used to imitate silence. The point here is that all three situations that are suitable for shin

---

16 We are grateful to Olivia Rohan, PhD candidate in the School of Applied Language and Intercultural Studies, Dublin City University, for sharing examples (23) and (24).
communicate different ‘feels’ or impressions of particular silences. The silence in (22) communicates the crisp and peaceful feel of a winter morning, while the silence in (23) involves the awkwardness of the situation. The silence in (24), on the other hand, yields a sense of unknown danger. The point is that the same onomatopoeia communicates a range of different impressions in different contexts, and it is not clear how a sound symbolism account explains such cases, especially where non-auditory sensory experiences come into play. Examples (25a) and (25b) illustrate this point:

(25a) [Description of bread]
shittori amafuwa ren-nyu pan
moist sweet-MIM milk-bread

‘moist, light, sweet milky bread’
(cookpad.com, 2015)

(25b) [Caption for a model photo]
amafuwa girlie short
sweet-MIM girlie short

‘sweet and airy girlie short-hair’
(beauty.hotpepper.jp, 2015)
In (25a) and (25b), the onomatopoeia *fuwa* is combined with *ama*, which is part of the adjective *amai* (sweet). In both cases, this semantically complex expression denotes something airy and sweet. It is interesting that not only are the two constituents of this compositional phrase from different sensory modes (taste and texture), but the composite expression is used to describe items in different modes: one in taste (bread), the other in vision (hairstyle). However, this is not particularly surprising, given Ward et al's (2006) argument that synaesthesia arises via a cross-modal channel rather than by activating two unimodal regions. Ward et al (2006) show that stimuli from particular domains can be ‘translated’ into others. Examples such as these show that onomatopoeia can be used to communicate representations of experience in different sensory domains.

As these examples show, from a communicative perspective, it does not matter whether the link between sound and meaning is arbitrary or not. If there is a link between sound and meaning, then a theory that appeals to such a link should be able to explain how the hearer would choose one ‘meaning’ over the others in a specific context. As it stands, no existing study on onomatopoeia seems to explain this. Relevance theory, on the other hand, enables us to explain this in terms of the relevance-guided comprehension heuristic: the hearer chooses one meaning over the others because of considerations of relevance. What is interesting is the fact that humans are capable of using such expressions as a tool to ‘show’ our perceptions to each other so that we can share impressions and feelings, which are quite often difficult to put into words. Some onomatopoeias might be more established as words and, of course, a link between
sound and meaning is established, as this stems from attempts by humans to ‘recreate’
their sensory experiences.

5. CONCLUSION

In this study, we have analysed onomatopoeia as a communicative phenomenon, and
argued that it falls on the showing - saying continuum, as suggested by Wharton (2009).
We have argued that many onomatopoeias have elements of both showing and saying,
and all provide direct evidence for the first layer of information that the communicator
intends to point out. As onomatopoeia often communicates extremely vague
impressions which are hard to render in purely propositional terms, it falls within the
We have shown that so-called sound symbolism, or the systematic (non-arbitrary)
relationship between sound and meaning, is a result of the communicator’s attempt to
recreate sensory experiences using whatever tools are available to him, and in
particular, by exploiting resemblances.

This communicative approach enables us to account for the subtle difference(s)
between closely-related onomatopoeias as well as the variety of ‘meanings’ a single
onomatopoeia can communicate.

This analysis, hopefully, sheds light on how showing can take place across different
modes and behaviours, which has not been explicitly investigated in the relevance-
theoretic literature before. This is made possible by treating onomatopoeia as a case of
showing in the modality of sound, while suggesting the possibility of treating mimetics as a case of cross-modal showing. This should pave the way for further investigation of the interface between verbal and non-verbal communication.

Acknowledgements

Words cannot express our gratitude to Diane Blakemore for her continuous help, support and generosity. She has been our inspiration and guidance. We are most grateful for her encouragement as well as the critical input that we will find insightful in years to come. We would also like to thank Deirdre Wilson for her support and comments on earlier versions of this paper, and to anonymous reviewers for their time and insightful comments.

References


Akita, K., 2013b. Onomatope/Otoshocho no kenkyu-shi.. In Shinohara, K., Uno, R. (Eds.), Onomatope Kenkyu no Shatei: Chikazuku Oto to Imi, 333-364, Hituji Shobo, Tokyo.


Kita, S., 2013. Giongo gitaigo to jihatsuteki miburi. In Shinohara, K., Uno, R. (Eds.), Onomatope Kenkyu no Shatei: Chikazuku Oto to Imi, 79-84, Hituji Shobo, Tokyo.


Toratani, K., 2013. Fukushiteki onomatope no tokushusei – tagisei/jishosei kara no kosatsu In Shinohara, K., Uno, R. (Eds.), Onomatope Kenkyu no Shatei: Chikazuku Oto to Imi, 85-99, Hituji Shobo, Tokyo.


