An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese[1]

M A R K   D I N G E M A N S E

Max Planck Institute for Psycholinguistics

K I M I   A K I T A

Nagoya University

Words and phrases may differ in the extent to which they are susceptible to prosodic foregrounding and expressive morphology: their expressiveness. They may also differ in the degree to which they are integrated in the morphosyntactic structure of the utterance: their grammatical integration. We describe an inverse relation that holds across widely varied languages, such that more expressiveness goes together with less grammatical integration, and vice versa. We review typological evidence for this inverse relation in ten spoken languages, then quantify and explain it using Japanese corpus data. We do this by tracking ideophones – vivid sensory words also known as mimetics or expressives – across different morphosyntactic contexts and measuring their expressiveness in terms of intonation, phonation and expressive morphology. We find that as expressiveness increases, grammatical integration decreases. Using gesture as a measure independent of the speech signal, we find that the most expressive ideophones are most likely to come together with iconic gestures. We argue that the ultimate cause is the encounter of two distinct and partly incommensurable modes of representation: the gradient, iconic, depictive system represented by ideophones and iconic gestures, and the discrete, arbitrary, descriptive system represented by ordinary words. The study shows how people combine modes of representation in speech and demonstrates the value of integrating description and depiction into the scientific vision of language.

KEYWORDS: depiction, expressiveness, ideophones, morphosyntax, semiotics, typology

---

1. INTRODUCTION

Words and phrases may differ in the extent to which they are susceptible to expressive features such as intonational foregrounding and expressive morphology: their EXPRESSIVENESS. They may also differ in the degree to which they are integrated in the morphosyntactic structure of the utterance: their GRAMMATICAL INTEGRATION. In this paper we describe an inverse relation that appears to hold across widely varied languages, such that more expressiveness goes together with less grammatical integration, and vice versa. We study this relation closely in ideophones, vivid sensory words found in many of the world's languages (Voeltz & Kilian-Hatz 2001, Dingemanse 2012). The findings shed light on the more general question of how language users cope with the challenge of combining distinct modes of representation in one modality.

Ideophones are a good category to study expressiveness and grammatical integration because they exhibit a particularly clear interaction between the two. They are often cast as prototypically expressive words that are only borderline linguistic: special in terms of sound patterns, suprasegmentals, syntax, and semantics (e.g. Kunene 1965, Zwicky & Pullum 1987). On a seemingly contrasting view, they may have mildly interesting morphosyntax, but are overall well integrated into broader linguistic systems (e.g. Newman 2001, Tsujimura & Deguchi 2007). We argue that these views can be unified in a typological-comparative approach that recognises that ideophones are often special words, set apart from other vocabulary in various ways, but that they can also enter into morphosyntactic constructions. We aim to explain why ideophones are typically expressive and free, and when they come to be more like ordinary words. We do this on the basis of descriptive data from a range of languages (Section 2), but most directly using corpus data from Japanese (Sections 3–5).

Japanese is well known for its extensive system of ideophones (Kita 1997, Hamano 1998, Akita 2009), vivid sensory words like ひらひら *ɕiɾaɕiɾa* 'fluttering', キビキビ *kibikibi* 'brisk, energetic', and ゴクン *gokɯn* 'gulping'. In Japanese linguistics, these words are generally known as 'giongo/gitaigo', the *gi*-component of which gives rise to the translated term 'mimetics'. Here we use the cover term IDEOPHONES, the most common cross-linguistic term for the phenomenon (Voeltz & Kilian-Hatz 2001). In Japanese as in other languages, ideophones can be defined as 'marked words that depict sensory imagery' (Dingemanse 2012): they stand out from ordinary words in terms of phonotactics, word structures, and morphosyntax, and they typically depict sensory imagery using various types of iconic mappings instead of representing it in arbitrary ways. Following Akita (2009), we define Japanese ideophones formally as those items that satisfy one of a limited number of morphophonological templates, such as reduplicative templates (e.g. $C_1'V_1C_2V_2$-$C_1V_1C_2V_2$) and suffixal templates (e.g. $C_1V_1C_2'V_2ri$). Akita (2009: 110) shows that these templates characterise 1643 out of 1652 items in a dictionary of Japanese ideophones. With a large and well-defined category of ideophones, Japanese is an excellent locus for a close study of expressiveness and grammatical integration.

As a first illustration of the phenomenon, compare the examples in (1). (Here and in subsequent numbered examples, bold indicates the ideophone in focus, upward arrows '↑' indicate intonational foregrounding, and vertical lines mark beginning and end of gesture Gx; in the English translation, the text in italics translates the ideophone, and the text following Gx provides a verbal description of the gesture.)

(1) (a) Sonoɯtʃi kawara-ga **gatʃagatʃag-gatʃaːt-to** otʃi-te
soon   tile-NOM   IDPH.PM1.VL.SR.VOICELESS-QUOT fall-CONJ
kɯ-rɯ.
come-NPST

                 |G1 |
'Then, the roofing tiles drop down on us *with a loud clattering noise*.'
G1: both hands loosely open, palms down, slightly moving up and down in front of the speaker's chest, synchronised with the production of the ideophone

                                                                                                      (D0007010069)

(b) Moː   boːhateː **girigiri-des**-ɯ.
already breakwater IDPH-COP.POL-NPST

'[The sea level] *was* already *almost reaching* the breakwater.'

                                                                                                       (D0007010082)

Both examples feature a disyllabic reduplicative ideophone: *gatʃagatʃa* 'clattering noise' and *girigiri* 'barely'. However, they differ in expressiveness, with the ideophone in (1a) showing partial multiplication, vowel lengthening, and voiceless phonation, but the ideophone in (1b) showing none of these signs of expressiveness. The ideophones also differ in grammatical integration: in (1a), the ideophone occurs in a quotative construction and is syntactically optional, whereas in (1b), the ideophone occurs in a nominal construction integrated into the predicate. This is the inverse relation between expressiveness and grammatical integration which we will be exploring in this study. Tellingly, the more expressive ideophone in (1a) comes with an iconic gesture, while the one in (1b) does not; a factor we will be able to use in our analysis.

We ask two research questions:

(i) Is there some systematic relation between expressiveness and grammatical integration?

(ii) If there is such a relation, what is the best explanation for it?

503

We use multimodal corpus data to answer these questions. We collect all utterances in which ideophones occur and keep track of (a) expressiveness, as measured by three features of the speech signal: intonational foregrounding, phonational foregrounding, and expressive morphology; (b) grammatical integration, in terms of construction types ranking from more to less morphosyntactic integration; and (c) iconic gestures, as a measure independent from the speech signal that can shed light on the reason for the special behaviour of ideophones. To answer question (i), we test whether measures of expressiveness covary with degree of morphosyntactic integration. To answer question (ii), we look at the nature of the relation between expressiveness and grammatical integration, and consider additional evidence from gestures co-occurring with ideophones.

2. EXPRESSIVENESS AND GRAMMATICAL INTEGRATION

Ideophones are often characterised as 'expressive' words. What exactly this means varies somewhat by author: in the context of ideophones, it is used in connection with affective content (Samarin 1970, Baba 2003, Potts 2007), experiential semantics (Klamer 2002, Blench 2013), and iconic form–meaning mappings (Diffloth 1980, Kakehi 1986, Tamori 1990). What unites these takes on expressiveness is that they can all be seen as pointing to the depictive nature of ideophones.

By 'depictive' we refer to one side of a well-known semiotic distinction between two modes of representation found in human communication: DESCRIPTION and DEPICTION (Clark & Gerrig 1990).[2] These two modes of representation differ in how they map form and meaning, how they are built, and how they are typically interpreted (Table 1). A common shorthand for the distinction is 'word' versus 'image', reflecting a traditional view of language as a system of arbitrary words fully in the descriptive mode, with the depictive method of communication at best playing a secondary role in the gestures and bodily aspects of 'paralanguage'. We argue that this traditional, exclusionary view may be profitably exchanged for a more inclusive account of the multiple semiotic resources that are available in everyday language use. Thus, where traditional accounts from Peirce (1955) to Hockett (1960) to Newmeyer (1992) have tended to equate spoken words with the descriptive mode of representation, we will provide evidence and arguments that speech can also be depictive.

---

[2] A third mode, indication, is represented by indexical signs but will not be our empirical focus. The three modes description, depiction and indication map onto Peirce's (1955) semiotic triad of symbol, icon, and index.

	Description (dieresis, digital, telling)	Depiction (mimesis, analog, showing)
Form–meaning mapping	Arbitrary	Iconic
Building blocks	Discrete symbols	Gradient markings
Interpretation	Decode to interpret	Imagine to interpret

Table 1
Two modes of representation and their prototypical features.

To foreshadow the argument, we will argue that ideophones typically are depictive representations in the verbal modality, and that recognising them as such explains certain widespread and otherwise unexpected regularities in the morphosyntactic typology of ideophones. To say that ideophones are depictions is to say that they are a form of mimesis rather than diegesis, that is they show rather than tell, they perform rather than merely inform. The proposal that ideophones are depictions is grounded in earlier work: they have been called vocal images (Westermann 1927), likened to gestures (Kunene 1965), and compared to cinematic imagery (Nuckolls 2000). What we contribute in this study is novel evidence from a multimodal corpus of Japanese, which brings into sharper relief when and how the special semiotic status of ideophones has repercussions for their morphosyntactic realisation.

While the semiotic distinction between description and depiction will prove important later on, the primary observables we are concerned with are in the domain of prosody and morphosyntax. We define the EXPRESSIVENESS of linguistic signs as the degree to which they are foregrounded as distinct from other items, for instance by special intonational or phonational features. This is in line with the established use of 'expressive' as a term that contrasts with 'plain', 'ordinary' or 'prosaic' (Fudge 1970, Diffloth 1980, Zwicky & Pullum 1987, Joseph 1994). A key notion here is foregrounding: 'the use of the devices of the language in such a way that this use itself attracts attention' (Havránek 1964: 10). Foregrounding is a semiotic means that may be used for different communicative ends, from indicating noteworthiness (Bolinger 1968) to expressing emotional content (Potts 2007), to framing something as a depictive performance (Nuckolls 1996). The fact that foregrounding is a flexible semiotic means is one of the reasons that the term 'expressiveness' has come to mean different things for different authors. For instance, emotional language and ideophones may both be foregrounded, but for different reasons: emotion words to index a particular affective stance, ideophones to index a particular mode of representation. Here we provide evidence that expressive foregrounding in ideophones is associated with a depictive mode of representation.

GRAMMATICAL INTEGRATION refers to the degree of integration in the morphosyntactic structure of the utterance. Grammatical integration can be measured in terms of linear position (peripheral items are less integrated), syntactic optionality (optional items are less integrated), and embedding in morphosyntactic structure (less deeply embedded items are less integrated). Across languages, the grammatical integration of ideophones tends to be low: they may be used alone as a complete utterance, and when they occur with a phrase they tend to appear at utterance edge in a loose appositional relation, where they may be grammatically optional (Watson 2001). However, ideophones do in fact partake in sentential structure to varying degrees, and focusing only on their prototypically free, expressive side would lead us to miss a number of observations on their morphosyntactic typology.

2.1 *Typological evidence*

The central question of this paper is motivated by a typological pattern that is relatively well attested in ideophone languages, though evidence for it has not, as far as we know, been brought together in one place before. It is typical for grammatical descriptions of ideophone systems to note the expressiveness of ideophones and, independently, their relative syntactic independence. Yet when grammars go into more detail, there are often hints of a more complex relation between both.

Table 2 lists ten languages from around the world for which grammatical descriptions provide some detail about these matters. In Bambara, a Mande language from Mali, ideophones normally occur at utterance edge, but 'non-final position of the ideophone causes it to lose its high intonation and its status as expressive adverb' (Dumestre 1998: 327). In Shona, a Bantu language spoken in southern Africa, ideophones often show expressive features of pitch and lengthening, but these tend to disappear when the ideophone is incorporated in a verb or derived into a noun (Fortune 1962). In Siwu, ideophones that are more deeply integrated in the morphosyntax lose their expressive features (Dingemanse 2013). In Somali, a Cushitic language of Somalia, ideophones are a subclass of nouns, and their noun-like character appears to be linked to them being less expressive (Dhoorre & Tosco 1998). In Semelai, an Aslian language of Malaysia, ideophones are 'never syntactically integrated' and 'usually . . . distinguished by an intonation break' from the surrounding material (Kruspe 2004: 399). In Jaminjung of Northern Australia, ideophonic coverbs are often intonationally foregrounded, particularly when they function as semi-independent predicates and to a lesser extent when they are incorporated in canonical complex predicates (Schultze-Berndt 2001). In Yucatec Maya of Mexico, when roots are instantiated as verbs, they are morphosyntactically integrated and combined with aspectual marking; but when instantiated as ideophones, the roots are characterised by syntactic independence, expressive morphology, and marked prosody (Le Guen 2012). In Pastaza Quechua of Ecuador, ideophones are often syntactically displaced or

isolated, in which case they are likely to be foregrounded intonationally, or they may occur in the unmarked preverbal position, in which case they are 'minimally performative because they are relatively assimilated into the intonational contours of their respective utterances' (Nuckolls 1996: 72). And finally, in Awetí, a Tupian language of Central Brazil, ideophones either occur as independent clauses or are embedded in a light verb construction. In the first case, they are 'always prosodically marked'; in the second, 'they may lose this feature and thus also their status as ideophones' (Reiter 2012: 576).

Language	Phylum	Macro-area	Source
Bambara	Mande	Africa	Dumestre 1998
Shona	Bantu	Africa	Fortune 1962
Siwu	Kwa	Africa	Dingemanse 2013
Somali	Cushitic	Africa	Dhoorre & Tosco 1998
Japanese	Japonic	Eurasia	Hamano 1998, *current study*
Semelai	Aslian	Eurasia	Kruspe 2004
Jaminjung	Mirndi	Australia	Schultze-Berndt 2001
Yucatec Maya	Mayan	North America	Le Guen 2012
Pastaza Quechua	Quechuan	South America	Nuckolls 1996
Awetí	Tupian	South America	Reiter 2012

Table 2
Ten languages of varied ancestry and geographical origins for which there is evidence of a relation between the expressiveness and grammatical integration of ideophones.

2.2 *Proposals and predictions*

The descriptions from the grammatical literature cited above all converge to suggest that some degree of expressiveness and syntactic freedom is the default case for ideophones, yet that things can also be turned around, with ideophones losing expressiveness when they are less syntactically free. Though these claims are suggestive, they are hampered by two problems: first, they are based on limited and in many cases unspecified data; second, they are unconnected observations without a unified explanation. What we aim to provide in this study is (i) an empirically grounded, quantitative investigation of the pattern in rich corpus data, and (ii) a unified account that explains the observations.

To structure the investigation, we note the following:

1. We observe an inverse relation between expressiveness and grammatical integration across languages. The more expressive an ideophone is (as measured by intonational foregrounding, phonational foregrounding, and expressive morphology), the less it is integrated in the morphosyntactic structure of the sentence.

2. We explain the inverse relation by reference to the semiotic nature of ideophones. In the canonical case, ideophones are performance-like depictions of sensory imagery, and to be realised and recognised as such, they have to be detached to some degree from the prosaic descriptive linguistic material that may surround them.
3. We predict that we can account for the expressiveness and morphosyntactic patterning of ideophones in a novel corpus of data (replicating the observation), and that we will find evidence of the depictive nature of ideophones when they are at their most expressive and free (supporting the explanation).

3. CURRENT STUDY: JAPANESE

Here we study expressiveness and grammatical integration using a corpus of Japanese. Corpus evidence is crucial for getting at expressive features of the speech signal, which would be hard to elicit and are best identified on the basis of recorded data available for repeated inspection. We use the NHK East Japan Great Earthquake Archives, one of the few available corpora of Japanese providing video as well as audio data of relatively informal, lively speech styles. All Japanese examples presented in this paper are taken from this corpus.

Previous work on Japanese has touched upon a link between expressiveness and grammatical integration. For instance, the syntactic embedding of ideophones has been linked to an 'iconicity continuum' (Hamano 1998) and syntactically independent ideophones were found in 'emotive' discourse (Baba 2003). The current paper builds on this work, but goes significantly further by providing detailed corpus evidence and an empirically grounded explanation for the observed inverse relation.

3.1 *Expressive features of ideophones in Japanese*

Expressive features, as defined above, are those features that help foreground ideophones as special words. Here we focus on three such features: intonational foregrounding, phonational foregrounding, and expressive morphology.

INTONATIONAL FOREGROUNDING is one of the most common expressive features associated with ideophones cross-linguistically (Childs 1994, Nuckolls 1996, Alpher 2001, Kruspe 2004, among many others). In many languages, ideophones are often produced at a markedly higher or lower pitch range. Sometimes there is also an intonational pause separating the ideophone from surrounding material. In Japanese, the prosodic peak of an utterance often coincides with the ideophone, already pointing to its special foregrounded status (Kita 1997: 395). Moreover, the pitch of the ideophone is often markedly distinct from the rest of the utterance. In (2) for instance, the ideophone *zabɯːn* 'splash' is produced in the upper part of the speaker's pitch range and after it, the speaker returns to the normal pitch register used for speaking (Figure 1).

(2) … ↑**zabɯːn**-**zabɯːn**↑-**tːe**-**i**-ɯ oto-wa wataʃi kikoe-te-ta-no.
 IDPH.VL.SR1-QUOT-say-NPST sound-TOP I hear-CONJ-PST-SFP

 '…I heard the sound like *splaash-splaash*.'
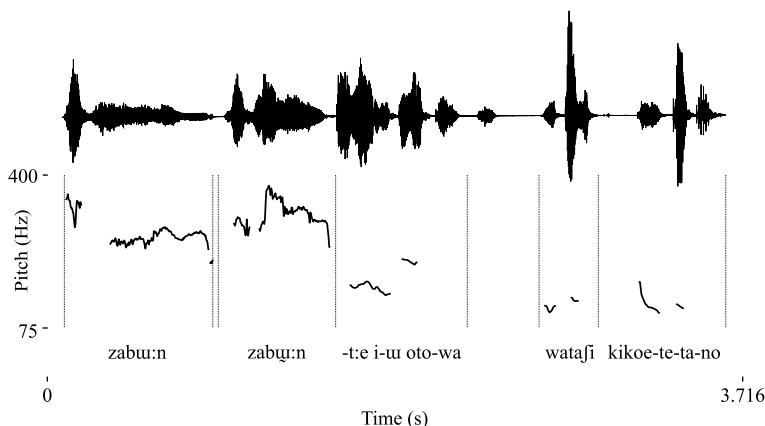
(D0007010147)

Figure 1
Waveform and pitch trace of (2), showing intonational foregrounding.

PHONATIONAL FOREGROUNDING is a term we introduce to capture the foregrounding of verbal material by means of marked departures from modal phonation. Though less frequently remarked upon than the intonational features of ideophones, special types of phonation such as breathy voice, growl, creaky voice, voicelessness, and whisper have been mentioned as a feature of ideophones across a range of languages (Childs 1994, Ameka 2001, Mihas 2012). In the Japanese earthquake corpus, we find phonational foregrounding of ideophones in the form of breathy voice, creaky voice, stiff voice, falsetto, voicelessness, or whisper. In (2) above, the second iteration of *zabɯ̰ːn* is produced with tense phonation close to creaky voice, which also affects the pitch. In (3a), *paːt-to* is pronounced as [pḁːtːo], with voiceless phonation. In (3b), *guːt-to* is pronounced as [gṵːtːo], with stiff voice, that is with the glottal opening narrower than normal.

509

(3)   (a)   Moː bikːɯɾi-ʃi-te   haʃit-te   waŋko-no   kɯsaɾi-o hodoi-te
             just IDPH-do-CONJ run-CONJ doggie-GEN chain-ACC untie-CONJ

           'O-ie-jo'-tːe   waŋko-ni   it-taɾa
           POL-house-SFP-QUOT doggie-DAT say-when

           paːt-to [pạːtːo]   hait-te.
           IDPH.VL.VOICELESS-QUOT enter-CONJ

           '[I] was astonished and ran, untying the doggie's chain and saying to the doggie, "House, boy," and then [it] entered [the house] *with a rush*.'

                                   (D0007010008)

    (b)   . . . moː akiɾaka-ni nagaɾe-ga **gɯːt-to** [gɯ̣ːtːo]   mata
             just obviously flow-NOM IDPH.VL.STIFF-QUOT again

           kotʃi-no   hoː-ni   hiɾogat-te . . .
           over.here-GEN direction-DAT spread-CONJ

           '. . . obviously, the flow spread *far and wide* over here again, and . . .'

                                     (D0007010092)

EXPRESSIVE MORPHOLOGY refers to special morphological processes applying commonly to ideophones and rarely to ordinary words, such as reduplication and lengthening (Zwicky & Pullum 1987).[3] Languages differ in the types of expressive morphology they make available. In Japanese, we find various types of stem repetition, partial multiplication, emphatic mora augmentation, vowel lengthening, and gemination (Hamano 1998, Nasu 2002, Akita 2009). For instance, the ideophone *don* 'bam' can undergo various processes of expressive morphology, from vowel lengthening (*doːn*) to partial multiplication (*dododon*) to stem repetition (*don-don-don*) (Akita 2009: 36).

Example (4a) illustrates vowel lengthening of the ideophone *gɯːt-to* 'rapidly', where the syllable is stretched around 300 ms longer than expected for normal speech (Figure 2). Two further examples of expressive morphology are in (4b), where *jɯkːɯɾi* 'slow' illustrates vowel lengthening and *gat(-to)* 'rattling' partial multiplication.[4] Finally, full repetition was illustrated in (2) above.

---

[3] Zwicky & Pullum (1987) have pointed to the special behaviour of ideophones to argue that expressive morphology 'constitutes a phenomenon that is not within the province of grammar as ordinarily understood' (1987: 338). Here, we aim to show that while ideophones are sometimes grammatically peripheral, they nonetheless partake in the morphosyntactic structure of utterances; and it is precisely the fact that they do so to varying degrees that allows us to investigate the inverse relation between expressiveness and grammatical integration.

[4] Additionally, *gagagagagagagat(-to)* is pronounced at a distinctly slower speech rate than the surrounding words, suggesting another possible type of prosodic foregrounding. Though we do not consider speech rate in a systematic way here, it may be another sign of the depictive use of speech (Childs 1994).

(4) (a) Mizɯ-ga ↑**gɯːt**↑**-to** ʃita-e sagat-te . . .
          water-NOM IDPH.VL-QUOT below-to go.down-CONJ

          'The water went down *rapidly*, and . . .'
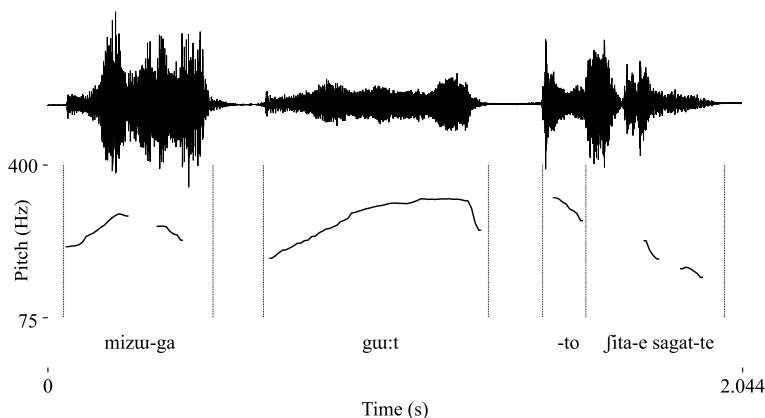
(D0007010164)

Figure 2
Waveform and pitch trace of (4a), showing expressive lengthening.

(b) . . . gareki-toka-ga ↑**jɯːkːɯri**↑ koː ɯgoi-te i-te
          debris-etc.-NOM IDPH.VL like.this move-CONJ be-CONJ
          soʃite mata **gagagagagagagat-to** ɕik-ɯ-no-ga
          and again IDPH.PM6-QUOT draw-NPST-NMLZ-NOM
          asa-made-ni maː ni-san-kai-de kika-na-i-gɯrai
          morning-until-DAT well 2–3-time-in suffice-NEG-NPST-degree
          at-ta-n-ʒa-na-i-kana.
          be-PST-NMLZ-COP.TOP-NEG-NPST-SFP

          '. . . things like debris moved *sloowly* and drew back *with a rattling sound*, which was [repeated] more than two or three times by the morning, I guess.'

(D0007010092)

   Intonational foregrounding, phonational foregrounding and expressive morphology are logically distinct, but they often occur together. Collectively, they contribute to the 'performative foregrounding' of ideophones (Nuckolls 1996): as we will see below, this foregrounding is done in the service of signalling a depictive mode of representation.

511

3.2 *Morphosyntax of Japanese ideophones*

The morphology and syntax of Japanese ideophones is a rich topic with a long history of research (Tamori 1984, Kita 1997, Hamano 1998, Kageyama 2007, Toratani 2007, Akita 2009, among many others). As illustrated in (5) below, Japanese ideophones are found in a range of morphosyntactic constructions (Tamori & Schourup 1999, Akita 2009, Toratani 2015). For each construction, the number of tokens observed in the corpus is given in parentheses.

(5) (a) Quotative (389 tokens)

Ie-ga   moː ↑**bakibaki**↑-**to** koware-te   it-te
house-NOM just IDPH-QUOT   break-CONJ go-CONJ
ʑibɯntatʃi-ga hait-te   i-rɯ   heja-ga   dondon
ourselves-NOM enter-CONJ be-NPST room-NOM one.after.another
katamɯi-te.
lean-CONJ

'Houses broke *with a cracking sound* one after another, and the room where we were leaned rapidly.'

(D0007010012)

(b) Collocational (155 tokens)

Donːa   kensa-o   ʃi-ta-no-ka-mo   **hakːiri**
what.kind checkup-ACC do-PST-NMLZ-Q-also IDPH
wakara-na-i-to-i-ɯ-koto-de . . .
know-NEG-NPST-QUOT-say-NPST-NMLZ-COP

'As [we] didn't even know *clearly* what kind of checkup [we] had conducted . . .'

(D0007010113)

(c) Noun-modifying 'say'-verbal (35 tokens)

. . . ʑibɯn-no atama-no naka-ni   aɾ-ɯ   tsɯnami-wa,
      self-GEN head-GEN inside-DAT be-NPST tsunami-TOP
ɕitaɕitaɕitaɕita-to-i-ɯ tsɯnami-nan-des-ɯ-jo.
IDPH-QUOT-say-NPST tsunami-COP-COP.POL-NPST-SFP

'. . . the tsunami I had in mind was something like *lap-lap-lap-lap*.'

(D0007010167)

(d) Predicative 'do'-verbal (58 tokens)

Sorede tʃot:o   kɯbi-o   daʃi-te   iki-o
then   a.little.bit neck-ACC put.out-CONJ breath-ACC

tsɯk-e-te   **hot-to-ʃi**-ta-n-des-ɯ-ne.
take-POS-CONJ IDPH-QUOT-do-PST-NMLZ-COP.POL-NPST-SFP

'Then, [I] could go out of [the water] and take a breath, and *got relieved*.'

(D0007010099)

(e) Predicative nominal (12 tokens)

… ʒuː-meːtorɯ-to-i-ɯ-no-wa   kono joŋ-kai-no
   10-meter-QUOT-say-NPST-NMLZ-TOP this 4-floor-GEN

jɯka-wa **girigiri-da**-so:-des-ɯ.
floor-TOP IDPH-COP-they.say-COP.POL-NPST

'…they say a ten-meter [tsunami] would *almost reach* this fourth floor.'

(D0007010115)

(f) Other (30 tokens)[5]

(i) Tabɯn   tateja-no   danːetsɯzai-mitai-na kanʒi-no,
     probably reactor.bldg-GEN insulation-like-COP feeling-GEN

     hirahira-to-ʃi-ta-joː-na   mono-ga Φɯt-te ki-ta.
     IDPH-QUOT-do-PST-like-COP thing-NOM fall-CONJ come-PST

     'Something *fluttering* that somehow looked like the insulation of the reactor building of the nuke rained on me.'

     (D0007010039; Noun-modifying 'do'-verbal)

(ii) Oki-no-hoː-wa   nantejɯːka, ɯzɯ-o
      distant.sea-GEN-direction-TOP how.to.say   swirl-ACC

      mai-te   **gotʃagotʃa-ni** nat-te
      roll-CONJ IDPH-COP   become-CONJ

      i-rɯ-n-des-ɯ-jo-ne.
      be-NPST-NMLZ-COP.POL-NPST-SFP-SFP

      '[I saw] the distant sea had, how to say, swirls and had been *jumbled up*.'

      (D0007010101; Nominal complement)

---

[5] This heterogeneous category includes the following low-frequency constructions (with the number of attested tokens in parentheses): Holophrastic (3), Predicative 'say'-verbal (2), Noun-modifying 'do'-verbal (5), Predicative 'become'-verbal (5), Nominal-intensificational (1), Noun-modifying nominal (3), and Nominal complement (11).

As the numbers of tokens suggest, five types of constructions are particularly common: Quotative, Collocational, 'say'-verbal (noun-modifying), 'do'-verbal, and Nominal (the latter two subsumed under the label 'Predicative'). These five types make available the two hierarchies for the degree of morphosyntactic integration of ideophones presented in (6). We describe the morphosyntactic basis for these rankings in the next section.

(6) *Grammatical integration of Japanese ideophones* (from less to more integrated)

    (a) Predicate integration hierarchy
        QUOTATIVE < COLLOCATIONAL < PREDICATIVE

    (b) Optionality hierarchy
        OPTIONAL < OBLIGATORY

3.2.1 *Predicate integration hierarchy*

The predicate integration hierarchy comprises four productive constructions that constitute a VP (accounting for 90% of ideophone tokens in the corpus): Quotative, Collocational, 'do'-verbal, and Nominal, the latter two united in being predicative. Ideophones in these constructions are integrated into the predicate to different degrees, as measured by obligatoriness and syntactic position.

The Quotative construction consists of an ideophone and the quotative particle *-to* or its colloquial counterpart *-te*. In the Quotative construction, ideophones are both syntactically and semantically separated from their host predicates (Toratani 2006, 2007; Akita & Usuki 2016). They are not syntactically obligatory elements; in (7) *bakibaki-to* 'with a cracking sound', *tʃiɾat-to* 'glancing' and *mɯkɯmɯkɯ-to* 'swelling up' can be left out without affecting the grammaticality.

(7) *Quotative*

    (a) Ie-ga   moː (↑**bakibaki**↑-**to**) koware-te
        house-NOM just IDPH-QUOT   break-CONJ
        it-te . . . (= (5a) above)
        go-CONJ
        'Houses broke (*with a cracking sound*) one after another, and . . .'

    (b) (↑**Tʃiɾat**↑-**to**) niwa-o   mi-taɾa joːsɯɾɯni kɯɾo-i   mizɯ-ga
        IDPH-QUOT yard-ACC look-if in.short black-NPST water-NOM
        doː-to   hait-te   ki-ta-n-des-ɯ.
        IDPH-QUOT enter-CONJ come-PST-NMLZ-COP.POL-NPST
        'When [I] took a *glance* at the yard, in short, a flood of black water came in.'

                                                (D0007010005)

(c) (**Mɯkɯmɯkɯ-to**) mak:ɯɾokɯ nat-te, nami-ga.
     IDPH-QUOT   deep.black   become-CONJ   wave-NOM

    'The wave got deep-black (*swelling up*).'

<div align="right">(D0007010205)</div>

Quotative ideophones frequently occur in preverbal (e.g. *bakibaki-to* 'with a cracking sound' preceding *kowaɾe-* 'break' in (7a)) and non-preverbal positions (e.g. *tʃiɾat-to* 'glancing' occurring away from *mi-* 'look' in (7b)). They collocate with a wide variety of predicates, as illustrated by *mɯkɯmɯkɯ-to* 'swelling up' in (7c), whose host predicate *mak:ɯɾokɯ naɾ-* 'get deep-black' is not directly predictable from its movement meaning. All these phenomena confirm the low integration of Quotative ideophones with the predicate.

    The Collocational construction has the ideophone appearing in close association with a verb, without a quotative marker. As illustrated in (8), ideophones in this construction appear to be omissible and may appear in different sentential positions; *hak:iɾi* 'clearly' in (8a) occurs preverbally, whereas *soɾosoɾo* 'walking gingerly' in (8b) occurs away from the predicate it modifies (*oɾi-* 'get down').

(8)   *Collocational*

    (a)   Don:a   kensa-o   ʃi-ta-no-ka-mo   (**hak:iɾi**)
         what.kind checkup-ACC do-PST-NMLZ-Q-also   IDPH

         wakara-na-i . . . (= (5b))
         know-NEG-NPST

         '[we] didn't even know (*clearly*) what kind of checkup [we] had conducted . . .'

    (b)   . . . (**soɾosoɾo**) jane-kara ori-te   it-te   gareki-o
               IDPH   roof-from get.down-CONJ go-CONJ debris-ACC

         aʃi-de   . . . tsɯt:sɯi-ta-n-des-ɯ-kedo . . .
         foot-by   poke-PST-NMLZ-COP.POL-NPST-but

         '. . . [I] (*gingerly*) went down off the roof and poked the debris by [my] foot, but . . .'

<div align="right">(D0007010106)</div>

However, ideophones in Collocational constructions form a tight unit with their host predicates (Akita & Usuki 2016), exhibiting a strong preference for preverbal occurrence and tending to have a close collocational relation with particular verbs. Working with a corpus of novels, Toratani (2006) reports that 81% of ideophones in Collocational constructions were found in the preverbal position, whereas only 48% of Quotative ideophones were preverbal. Moreover, the Collocational construction is far less likely than the Quotative construction to allow atypical collocations, such as (7c) above (?*mɯkɯmɯkɯ mak:ɯɾokɯ naɾ-* 'get deep-black swelling up'). These syntactic and collocational facts are signs that ideophones

occurring in the Collocational construction are more morphosyntactically integrated than in the Quotative construction.

The two Predicative constructions – 'do'-verbal and Nominal – are more tightly integrated with the predicate. The 'do'-verbal construction consists of an ideophone and the dummy verb *su-* 'do' (Tsujimura 2005, Kageyama 2007), and the Nominal construction consists of an ideophone and a copula (Kita 1997, Toratani 2015). These ideophonic constructions are not syntactically optional, as illustrated in (9) and (10).

(9) *Predicative 'do'-verbal*

    (a) ... iki-o   tsɯk-e-te
           breath-ACC take-POS-CONJ

        *(**hot-to-ʃi**-ta-n-des-ɯ-ne). (= (5d) above)
         IDPH-QUOT-do-PST-NMLZ-COP.POL-NPST-SFP

        '[I] could take a breath, and *(*got relieved*).'

    (b) Soːʃitaɾa moː   kɯɾɯma-ni not-te   i-ɾɯ   çito-ga
        then   already car-DAT ride-CONJ be-NPST person-NOM

        *(**ʃoroʃoro-ʃi**-te i-ɾɯ-wake).
         IDPH-do-CONJ be-NPST-SFP

        'Then, people who had already got in a car *(were *wandering around*).'

                                       (D0007010010)

(10) *Predicative nominal*

    (a) ... kono joŋ-kai-no   jɯka-wa
           this 4-floor-GEN floor-TOP

        *(**girigiri-da**-soː-des-ɯ). (= (5e) above)
         IDPH-COP-they.say-COP.POL-NPST

        '... they say [the tsunami] would *(*almost reach*) this fourth floor.'

    (b) ... minːa moː   tatemono-toka-de *(**gɯʒagɯʒa-de**),
           all   already bldg-etc.-with   IDPH-COP

        toːɾ-e-nakɯ-nat-te   i-te ...
        pass-POS-NEG-become-CONJ be-and

        '... [the street] *(*was* already all *wrecked*) with buildings, and were impossible to go ...'

                                       (D0007010105)

These data are the basis for the predicate integration hierarchy for Japanese ideophones in (6a). The Quotative construction is least integrated in that it is syntactically optional and freely occurs in different sentential positions. The Collocational construction is more integrated into the predicate in terms of its positional and collocational preference. The 'do'-verbal and the Nominal constructions are part of the predicate and constitute the indispensable part of

a sentence. This three-way hierarchy allows us to examine the inverse relation between grammatical integration and expressiveness in a gradual fashion for the majority of ideophone tokens in the corpus.

3.2.2 *The optionality hierarchy*

A second, simpler way of measuring morphosyntactic integration is by syntactic optionality only, a measure that can be applied to all ideophones in the corpus regardless of construction type.[6] Given this criterion, the ideophonic constructions in Japanese can be divided into two groups, as in (11).

(11) *Syntactic optionality of ideophonic constructions*

    (a) OBLIGATORY
Predicative 'do'-verbal, Predicative 'become'-verbal, Nominal complement, Predicative nominal

    (b) OPTIONAL
Holophrastic, Quotative, Collocational, Noun-modifying 'say'-verbal, Noun-modifying 'do'-verbal, Noun-modifying nominal

The constructions in (11a) are syntactically obligatory, as they constitute predicate complexes. In contrast, the constructions in (11b) are syntactically optional, as they either stand alone or modify predicates or nouns as additional elements. We assume that obligatory elements are more deeply grammatically integrated into the sentence structure than optional elements. This simple dichotomy can be used in addition to the three-way predicate integration hierarchy to investigate the relation between grammatical integration and expressiveness.

3.3 *Gesture*

So far we have considered only speech, distinguishing expressive features and construction types to be tracked for each ideophone token in our corpus of Japanese narratives. Yet language use is fundamentally multimodal, and utterances often combine speech and gesture. Gesture provides an additional source of evidence that may help elucidate the mechanism underlying the interaction we observe. Gestures come in many types (Kendon 2004), but given our focus on depiction, we will primarily be interested in ICONIC GESTURES: manual movements whose shape depicts aspects of meaning by means of perceptual analogies. Consider the following examples from the corpus:

---

[6] We thank an anonymous *JL* referee for suggesting to apply a single criterion across the board as another measure of grammatical integration.

(12) . . . sonomama ↑**ba:n**↑-**t:e**-i-ɯ   katatʃi-de mizɯ-ga
        as.it.is   IDPH.VL-QUOT-say-NPST form-COP water-NOM
                    |G1 |

hait-te-ki-ta-node   takasa-t:e-i-ɯ-no-wa
enter-CONJ-come-PST-because height-QUOT-say-NPST-NMLZ-TOP
dakara   wakara-nakat-ta.
therefore know-NEG-PST
|G2| |G3|
'. . . as the water came in as it was in the form like *baang*, [I] didn't see [its]
height.'

(D0007010111)

G1: iconic, both open palms moving toward the speaker's face (Figure 3)
G2: iconic, both open palms moving toward the speaker's face (lighter than
G1)
G3: deictic, pointing towards the speaker's back

Figure 3
(Colour online) Gesture 1 from (12), accompanying *ba:n* 'incoming wave'.

(13) Mo: ɯɴ dakaɾa miɴːa   ↑**zoɾozoɾoːt**↑-**te**   koko haʃit-te.
        just yeah y'know everyone IDPH.PM2.VL-QUOT here run-CONJ
                                  |G1 | |G2 |
'Yeah, everyone ran through here *one after another in line*.'
G1: iconic, both arms swinging from the speaker's right to left (Figure 4)
G2: iconic, both arms swinging from the speaker's right to left (briefer and
less intense than G1)

(D0007010035)

Figure 4
(Colour online) Gesture 1 from (13), accompanying *zorozoroːt* 'one after another in line'.

In example (12) above, the ideophone *baːn* is produced with intonational foregrounding and expressive lengthening and occurs in the noun-modifying 'say'-verbal construction. It comes time-aligned with an iconic gesture depicting the movement of the enormous wave of water approaching (Figure 3). In example (13), the ideophone *zorot*(*-te*) 'one after another in line' is likewise produced with intonational foregrounding and expressive morphology and occurs in the Quotative construction. It comes time-aligned with an iconic gesture, in which the movement of both arms depicts the 'one after another in line' meaning evoked by the ideophone (Figure 4).

The relation between ideophones and iconic gestures is not exclusive: iconic gestures may co-occur with other elements as well, as with the verb in (12). However, earlier work has shown that ideophones are much more likely to co-occur with iconic gestures than verbs are: in a Japanese corpus of cartoon retellings, as many as 94% of ideophones co-occurred with an iconic gesture, whereas this held for only 40% of verbs in a matched sample from the same corpus (Kita 1997; see also Son 2010). This shows that there is a privileged relation between ideophones and iconic gestures, probably reflecting their shared nature as depictions of sensory imagery (Kunene 1965).

Gestures occurring with ideophones are closely related to them in two key ways: (i) they coincide temporally, with gesture and ideophone starting around the same time and gestures being repeated if ideophones are repeated; and (ii) the information they reveal is closely related to the meaning of the ideophone, with the visual iconicity of manual gestures joining the vocal iconicity of ideophonic speech (Nuckolls 2000). Studies of co-speech gesture often assume that speech and gesture play complementary roles, with speech supplying the relatively abstract verbal content and iconic gestures supplying more imagistic, gradient representations (McNeill 1992; Özyürek 2014; Goldin-Meadow, published online 1 July 2016; Kelly, in press). But when it comes to ideophones, speech and gesture are not loosely aligned and complementary, but tightly coupled and alike in mode of representation: they perform the same role of depicting sensory imagery, albeit in different modalities and therefore also with different affordances for iconicity (Dingemanse 2013).

All this implies that we can use evidence from co-occurring iconic gestures to shed light on the mechanism underlying the relation between expressiveness and integration investigated here. Essentially, iconic gesture enables us to link expressiveness and grammatical integration to the distinction between descriptive and depictive modes of representation. If ideophones are at their most expressive and least integrated when they are most like depictions of sensory imagery, then they should be especially likely to come together with gestures in exactly those conditions.

4. MATERIALS AND METHOD

The NHK East Japan Great Earthquake Archives corpus consists of the videos and transcripts of 214 interviews with victims and rescuers in the great earthquake which hit East Japan on 11 March 2011.[7] The length of the interviews ranges from about five to fifteen minutes, and the whole database contains 10,657 utterances. The recordings usually start with edited summary, followed by a free-form narrative; we consider only ideophones in the narratives. Within the narratives, we find 692 ideophone tokens (203 types).

We coded all 692 ideophone tokens in the corpus for construction type and syntactic optionality, as well as for the occurrence of intonational foregrounding, phonational foregrounding, and expressive morphology. Additionally, we coded the ideophones for co-occurring gestures. There were 549 cases in our data where the speaker of the ideophone is visible in the frame. In 284 of these (52%), the ideophone occurred with a time-aligned gesture. Of these gestures, 94% were iconic, 4% were pointing gestures, and a handful were mixed or beat-like movements (see Kendon 2004 for a discussion of these categories). This confirms the strong association between ideophones and iconic gestures found in earlier work (Kita 1997, Dingemanse 2013).

Whereas measures of expressive morphology and syntactic integration are unambiguous – based on well-defined rules of ideophonic morphophonology and morphosyntax (Hamano 1998, Akita 2009) – assessing the presence or absence of intonational foregrounding, phonational foregrounding, and iconic gestures can be more subjective. To ensure coding reliability, 10% of the data (70 cases) were also coded by an independent coder. Concordance rates and Cohen's κ show that coding reliability is good to very good (Table 3). Measures like Cohen's κ assume equiprobability and penalise skewed data values – the 'high agreement, low consistency' paradox (Feinstein & Cicchetti 1990). For intonational foregrounding, our interpretation of the κ value is 'good' because the two values of intonational foregrounding (yes and no) are not equiprobable in our data while the concordance rate is still 74.29%.

---

[7] The present study is based on the interviews that were available on 28 February 2013. Since then, the number of interviews in the NHK archive has grown.

	Concordance rate and Cohen's κ
Intonational foregrounding	74.29%, $\kappa = .47$, $p < .001$
Phonational foregrounding	87.14%, $\kappa = .71$, $p < .001$
Iconic gestures	80.00%, $\kappa = .68$, $p < .001$

Table 3
Coding reliability.

5. FINDINGS

We hypothesised that the degree of expressiveness of ideophones is inversely related to their grammatical integration. On this hypothesis, the more integrated an ideophone is, the less expressive features it should show. If the hypothesis is false, the degree of expressiveness of ideophones should bear no relation to their grammatical behaviour.

5.1 *Descriptive statistics and correlations*

Of the 692 ideophone tokens, 625 can be categorised as Quotative (389), Collocational (155) or Predicative (81), confirming that ideophones on the whole tend towards grammatical independence. However, here we are interested in how ideophones behave across the cline of grammatical integration, so we track expressive features for each of the three morphosyntactic contexts.

Figure 5 shows the presence and absence of expressive features across the Predicate integration hierarchy, Quotative < Collocational < Predicative in the order of increasing integration. The distribution is as predicted by the main hypothesis: a disproportionate amount of ideophones with expressive features occur in the Quotative construction, less so in the Collocational construction, and least of all in Predicative constructions. The three measures of expressiveness all pattern in broadly similar ways, though intonational foregrounding is most frequently attested and phonational foregrounding is much less frequent overall.

Statistical tests confirm the distributional evidence shown in Figure 5: morphosyntactic integration is negatively correlated with intonational foregrounding (Pearson's $r = -0.31$, $p < .0001$), expressive morphology ($r = -0.38$, $p < .0001$), and phonational foregrounding ($r = -0.33$, $p < .0001$), all $n = 625$, all p-values Bonferroni-corrected for multiple comparisons. In short, there is an inverse relation between each of the individual expressive features and the degree of morphosyntactic integration of ideophones.
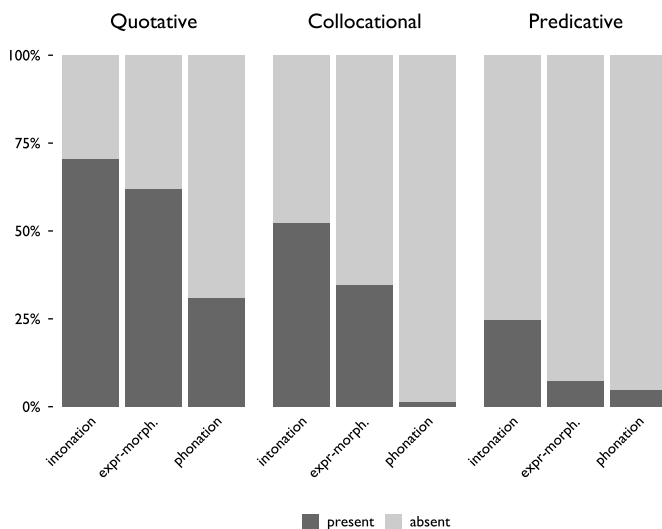
Figure 5
Expressive features of ideophones by morphosyntactic integration ($n = 625$).

As each of the expressive features can occur on their own, they are logically independent. Despite this, they are highly correlated with each other, as expected on the hypothesis that they are all indexes of the depictive mode of representation. We see positive correlations between intonational foregrounding and expressive morphology ($r = 0.46$, $p < .0001$), intonational and phonational foregrounding ($r = 0.22$, $p < .0001$), phonational foregrounding and expressive morphology ($r = 0.36$, $p < .0001$), all $n = 625$, all p-values Bonferroni-corrected for multiple comparisons.

To further study the relation between expressiveness and grammatical integration, we compute a cumulative measure of expressiveness ranging from 0 (no expressive feature) to 3 (three expressive features). Figure 6 shows the grammatical integration and cumulative expressiveness for ideophone tokens in the corpus. Ideophone tokens with the highest degree of expressiveness are found almost exclusively in the Quotative construction; ideophone tokens with two or one expressive feature are found predominantly in the Quotative and Collocational constructions. In contrast, the majority of ideophones in Predicative constructions show no expressive features at all.

The skewness of the distribution is borne out in a statistical test, which finds a strong negative correlation between cumulative expressiveness and morphosyntactic integration (Spearman's $\rho = -0.45$, $p < .0001$, $n = 625$). We use a Spearman ranked correlation here instead of Pearson's rho because the comparison is now between two ordinal variables ($0 < 1 < 2 < 3$ for cumulative expressiveness; Quotative $<$ Collocational $<$ Predicative for grammatical integration). There are no differences in significance when running a Pearson correlation.
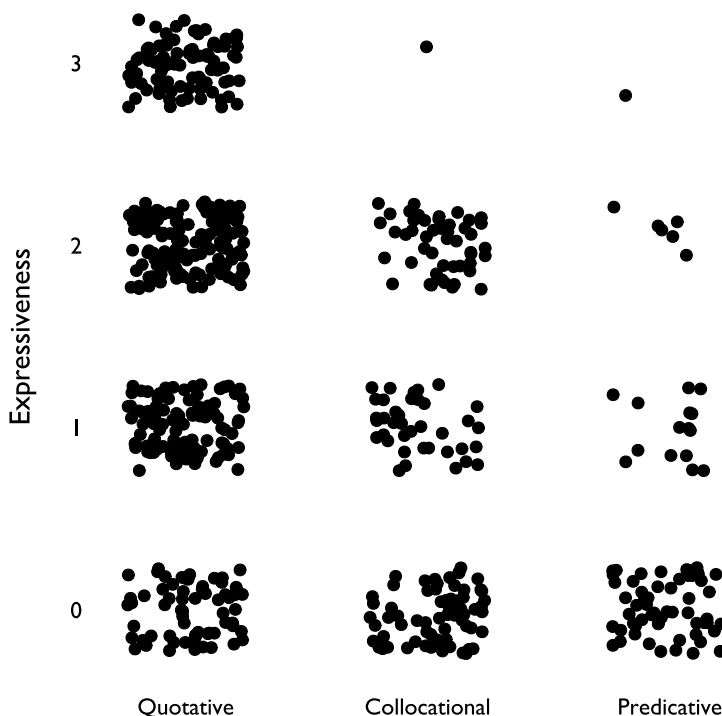
Figure 6
Cumulative expressiveness and morphosyntactic integration for ideophones in the corpus
($n = 625$), showing that higher expressiveness correlates with lower integration.

5.2 *Linear mixed effects modelling*

As we work with corpus data, the number of cases in each condition is not balanced, and there are multiple possible dependencies between observations: some utterances in the corpus feature multiple ideophones, and some utterances are produced by the same speaker, so not all observations are independent and they may be affected by individual differences in expressive speech. We use mixed effects modelling (Bates, Mächler, Bolker & Walker 2015) in R (R Core Team 2015) to find out whether cumulative expressiveness remains a significant predictor of morphosyntactic integration when we control for these dependencies.

We constructed a mixed effects model of morphosyntactic integration as a function of expressiveness. The fixed effect factor is expressiveness. We include utterance as a random effect and allow expressiveness to vary with random slope by narrative (each narrative has a unique narrator, so this is a proxy for individual differences). The model achieves a fit that is significantly better than a null model with no fixed effect ($\chi^2(1) = 91.33$, $p < .0001$, log likelihood difference = 45.66). Thus, mixed effects modelling confirms that higher expressiveness is related

to lower grammatical integration even when we control for dependencies and imbalances in the corpus data. If we make the simplifying assumption that the distance between categories is the same at every scale point, we can interpret the model estimate as telling us that every step up in expressiveness makes it 29% more likely that an ideophone is realised with lower grammatical integration.

So far we have focused on grammatical integration as measured by the predicate integration hierarchy, our most fine-grained view of the phenomenon. The results are strongly similar for integration measured in terms of syntactic optionality. There is a strong correlation between cumulative expressiveness and syntactic optionality (Spearman's $\rho = 0.27$, $p < .0001$, $n = 692$). A mixed effects model testing how cumulative expressiveness affects optionality (with the same random effects structure as above) achieves a fit that is significantly better than a null model with no fixed effect ($\chi^2(1) = 32.31$, $p < .0001$, log likelihood difference = 16.15). So higher cumulative expressiveness is consistently correlated with lower grammatical integration whichever measure of integration we use.

The results provide a detailed view of the relative expressiveness of ideophones across morphosyntactic contexts. The three expressive features pattern in the same way: they are most common in the most free construction types and least common (in fact mostly absent) in the least free ones. Correlational evidence and mixed effects modelling converge to confirm the main hypothesis that expressiveness is inversely related to morphosyntactic integration.

5.3 *Evidence from gesture*

So far, all our measures have been from the speech signal and its morphosyntactic structure. As we have argued above, evidence from gesture may serve to elucidate the underlying mechanism: if expressive features are indicative of the depictive mode of representation, this should be reflected in the patterning of iconic gestures, which also inhabit this mode of representation. So we expect iconic gestures to be particularly common when ideophones are at their most expressive.

Like Figure 6 above, Figure 7 below shows all ideophones according to the predication hierarchy, but now coded to show the co-occurrence of iconic gestures (which are observed for 242 tokens). Gestures accompanying ideophones are most abundant in the Quotative construction (58% of visible tokens), less so in the Collocational construction (41%), and least in the Predicative constructions (18%). There is also a clear relation to cumulative expressiveness: gestures are most abundant for ideophones with three expressive features (85% of visible tokens), less so for tokens with two expressive features (72%), less again for one expressive feature (48%), and least of all for ideophones without expressive features (only 12%). A statistical test confirms that there is a strong negative correlation between the occurrence of gesture and morphosyntactic integration (Pearson's $r = -0.26$, $p < .0001$, $n = 492$), and a strong positive correlation between gesture and cumulative expressiveness (Pearson's $r = 0.55$, $p < .0001$, $n = 492$, all p's Bonferroni-corrected).
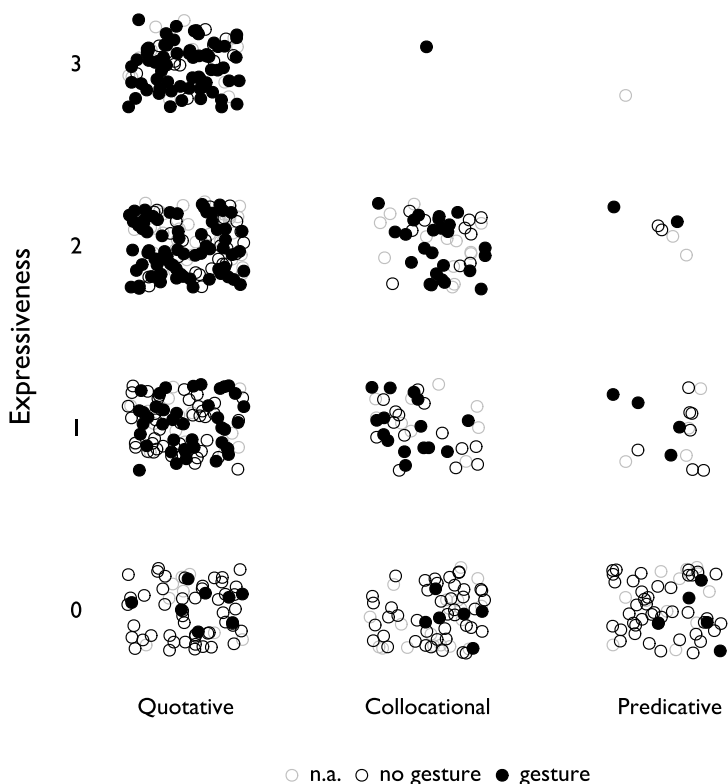
Figure 7
Gesture, expressiveness and integration for ideophones in the corpus ($n = 625$), showing
gestures ($n = 242$) tend to co-occur with ideophones when they are most expressive and
least integrated.

In sum, co-occurring gestures are a strong predictor of the morphosyntactic integration of ideophones, and are highly correlated with expressive features of the speech signal. This confirms the main hypothesis and supports the proposed mechanism by showing that the inverse relation between expressiveness and grammatical integration is directly connected to the depictive nature of ideophones.

6. DISCUSSION

Our results accomplish three things: (i) they provide empirical grounding for the relation between expressiveness and grammatical integration; (ii) they provide a more detailed view of the interaction, teasing apart three clearly defined expressive features and showing for each of them how they vary across three morphosyntactic contexts; and (iii) they use evidence from gesture to link the

inverse relation between expressiveness and integration more directly to the depictive nature of ideophones. We are now in a position to move from observation to explanation.

6.1 *Explanation*

Why do expressiveness and morphosyntactic freedom go together so naturally in ideophones? Conversely, why do ideophones tend to lose their expressiveness when they are more deeply integrated in the utterance? Both expressiveness and a lack of morphosyntactic integration are often cited as typical features of ideophones, so to cite their co-occurrence as explaining the relation would border on tautology. Instead we propose that the expressive and free nature of ideophones has to do with a fundamental difference in mode of representation: they are depictions as opposed to descriptions.

The interaction between expressiveness and grammatical integration reflects the encounter of two partly incommensurable methods of communication: the discrete, arbitrary, descriptive system represented by ordinary words, and the gradient, iconic, depictive system embodied by ideophones. These two methods place different requirements on the material use of speech: in description, discrete segments like phonemes and morphemes are combined, integrated and linearised into ordinary utterances; in depiction, speech is used in a more gradient way to suggest meaning by means of iconic form–meaning mappings. Yet both are inevitably part of the same single linearly unfolding speech stream, which leads to a challenge akin to the linearisation problem in psycholinguistics (Levelt 1981).

We propose this challenge is met by the relation between expressiveness and syntactic independence we see in ideophones. The expressive features foreground the ideophone, drawing attention to the word as a performance and inviting listeners to imagine what it is like to perceive what it depicts. This performance requires a 'stage' where the vivid depiction of sensory imagery in speech is the rule rather than the exception. This is why ideophones thrive when they are free, and why they behave more like plain words when bound within morphosyntax. In the words of Daniel Kunene,

> The ideophone stands aloof from the connecting tissues, the sinews and ligaments that flesh out the basic components of speech into a morphological, grammatical and syntactical system. By thus isolating itself, it, so to speak, climbs the stage to become an act, thus removing itself from the run-of-the-mill narrative surrounding it. (Kunene 2001: 190)

The contribution of the present study has been to put this explanation of ideophone morphosyntax on firm empirical footing. We first tested the relation between expressive features of the speech signal and morphosyntactic freedom and found an inverse relation: more expressiveness goes together with less integration. We then tested whether freedom and expressive features are signs of the depictive nature of ideophones using a measure distinct from the speech signal: gesture. We found that the expressive features OF spoken ideophones

strongly correlate with the iconic gestures coming WITH them, strengthening the case for the analysis of ideophones as depictive performances when at their most expressive and free.

Although the primary data in this study have come from Japanese, we have brought together preliminary evidence for strongly similar patterns in a wide range of unrelated languages (Table 2). These descriptive facts of ideophone typology receive a unified explanation in the conceptual framework outlined here.

6.2 *Predictions*

Besides explaining a good number of empirical observations, our account generates testable predictions in the areas of language typology, language processing, and language change.

With regard to language typology, our account predicts the absence of languages in which highly integrated ideophones tend to be more expressive than loosely integrated ideophones. Such languages would falsify the account in its current formulation. Instead we expect that ideophones in any language will show the same inverse relation between expressiveness and grammatical integration. Relatedly, our account predicts that features of expressiveness and measures of morphosyntactic integration of ideophones may differ as a function of the typological profile of languages. For instance, some languages may allow ideophones to occur on their own, others may always require the presence of some grammatical marker, and yet others may feature ideophones that are even more bound, and each of these will have consequences for the (relative) expressiveness of ideophones in that context. Further, our results provide the first quantitative evidence of phonational foregrounding and predict this is likely more widespread than reported so far.

The regularities we have described here provide people with a set of heuristics that can guide the production and interpretation of verbal material. From this follows a prediction that ideophones produced with more expressive features should invite more iconic interpretations. Indeed, recent experimental work shows that prosody (including intonation and duration) is an important contributor to sound-symbolic effects (Nygaard, Herold & Namy 2009, Dingemanse, Schuerman, Reinisch, Tufvesson & Mitterer 2016). Another prediction is that newly created ideophones, which may be based on new or existing verbal material, should be maximally expressive and minimally integrated in order to be recognised as depictions rather than descriptions. Both predictions can be experimentally tested and point to promising directions for future research on ideophones and vocal iconicity.

Further predictions arise in the domain of language change. If new ideophones must be maximally expressive, it follows that less expressive ideophones are likely older and more conventionalised. So the inverse relation between expressiveness and grammatical integration provides a pathway for deideophonisation. Over time, ideophones may assimilate to ordinary vocabulary. Our account predicts

that this assimilation will manifest itself as a gradual loss of expressiveness and a gradual increase in morphosyntactic integration. Known processes of reduction and conventionalisation (Bybee 2007) suggest that this is especially likely to occur with more frequently used ideophones. The overall picture that emerges is one of ideophone inventories as in a state of flux, with expressiveness and grammatical integration pulling in different directions.

6.3 *Generalisations*

The inverse relation between expressiveness and grammatical integration is not limited to ideophones. It applies to any situation in which language users combine descriptive and depictive content.

Take quotations. These can fruitfully be analysed as depictions or 'demonstrations' (Clark & Gerrig 1990, Davidson 2015) in which speakers attempt to depict selected aspects of some original (speech) behaviour. Like ideophones, quotations come in a range of construction types that differ in their morphosyntactic integration (De Vries 2008, Güldemann 2008). There is also evidence that this difference goes along with prosodic cues akin to the expressive features we have discussed for ideophones; for instance, in direct quotation (but not indirect quotation), the quoted speech tends to be set off prosodically from the surrounding material by means of pauses and intonational foregrounding (Güldemann 2008: 222–223). Furthermore, direct quotations often include manual gestures and other visual signals that help signal the depictive mode of representation (Sidnell 2006). The picture proves to be exactly parallel: the more expressive features a quotation shows, the more likely it is to be morphosyntactically free; and both freedom and expressiveness are signs of its status as a depiction.

Ideophones and verbal quotations are special in being vocal depictions occurring within the linearly unfolding stream of speech. The inverse relation between expressiveness and integration directly follows from this fact: within the confines of the modality of speech, the main way to differentiate depiction from description is to exploit the temporal and material properties of the speech stream. This is why depictions in speech often occur at utterance edge, clearly distinguished from the adjacent descriptive material.

Speech is of course only one aspect of the rich reality of face-to-face interaction, which often combines different articulators and semiotic resources into composite utterances (Slama-Cazacu 1976, Enfield 2009). From this perspective, the relation between expressiveness and integration captures a more fundamental fact about the production and interpretation of composite utterances in communication. If we use different modes of representation together, there must be cues or meta-communicative signals (Bateson 1955) that enable language users to distinguish these modes and parse and interpret the material accordingly. Sometimes such cues may be given by the fact that the semiotic resources are materially distinct and used according to their most natural affordances. This is what gives rise to default inferences linking manual gesture to depiction and

speech to description (Goldin-Meadow, published online 1 July 2016). When this is not the case, cues must be realised in material aspects of the signals. In this study we have investigated various expressive features of speech: modifications of the signal that serve to cue a distinction in mode of representation. In other modalities we can expect analogous modifications or semiotic resources to serve as meta-communicative signals. For example, shifts in gaze patterns and body posture can play a similar role in demarcating depictions in multimodal discourse (Sidnell 2006) and in the phenomenon of role shift in sign language (Davidson 2015).

We have focused here on ideophones as they provide a particularly clear view of the inverse relation between expressiveness and grammatical integration. Exclamations and interjections form another of group of linguistic signs combining high expressiveness and low integration. These signs are primarily indexical (Kockelman 2010) as opposed to depictive or descriptive. Therefore, their expressiveness and grammatical freedom across languages may be explained by a generalised version of our account: it is a way to meet the challenge of combining distinct modes of representation in a linear speech signal.

7. CONCLUSIONS

Descriptions of ideophones have often been torn between casting them as extralinguistic exotics or stressing their integration in linguistic systems. Here we have shown how these seemingly contrastive views can be reconciled in a more comprehensive view of the morphosyntactic typology of ideophones – and how in the process we can shed light on a more general aspect of human language. Intonational foregrounding, phonational foregrounding, and expressive morphology commonly accompany ideophones across languages. They are not merely incidental features correlating with ideophone use, but serve the semiotic function of foregrounding some stretch of the speech signal as a depiction as opposed to a description.

The inverse relation between expressiveness and grammatical integration we have described here shows how people may exploit material features of the speech signal to conjoin and contrast different modes of representation. Too often, the distinction between description and depiction has been used as a demarcation line separating word from image and language from paralanguage. Here we have shown that it runs through much of our everyday language use, and that we can only begin to understand language in all of its aspects when we consider description and depiction together.

REFERENCES

Akita, Kimi. 2009. *A grammar of sound-symbolic words in Japanese: Theoretical approaches to iconic and lexical properties of Japanese mimetics*. Ph.D. dissertation, Kobe University. http://www.lib.kobe-u.ac.jp/handle_gakui/D1004724.
Akita, Kimi & Takeshi Usuki. 2016. A constructional account of the 'optional' quotative marking on Japanese mimetics. *Journal of Linguistics* 52.2, 245–275.

Alpher, Barry. 2001. Ideophones in interaction with intonation and the expression of new information in some indigenous languages of Australia. In Voeltz & Kilian-Hatz (eds.), 9–14.

Ameka, Felix K. 2001. Ideophones and the nature of the adjective word class in Ewe. In Voeltz & Kilian-Hatz (eds.), 25–48.

Baba, Junko. 2003. Pragmatic function of Japanese mimetics in the spoken discourse of varying emotive intensity levels. *Journal of Pragmatics* 35, 1861–1889.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1, 1–48.

Bateson, Gregory. 1955. A theory of play and fantasy. *Psychiatric Research Reports* 2.39, 39–51.

Blench, Roger. 2013. Mwaghavul expressives. In Henry Tourneux (ed.), *Chadic linguistics 8*, 53–75. Köln: Rüdiger Köppe.

Bolinger, Dwight L. 1968. *Aspects of language*. New York: Harcourt, Brace and World.

Bybee, Joan L. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Childs, G. Tucker. 1994. African ideophones. In Hinton et al. (eds.), 178–204.

Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66.4, 764–805.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38.6, 477–520.

De Vries, Mark. 2008. The representation of language within language: A syntactico-pragmatic typology of direct speech. *Studia Linguistica* 62.1, 39–77.

Dhoorre, Cabdulqaadir Salaad & Mauro Tosco. 1998. 111 Somali ideophones. *Journal of African Cultural Studies* 11.2, 125–156.

Diffloth, Gérard. 1980. Expressive phonology and prosaic phonology in Mon-Khmer. In Theraphan L. Thongkum (ed.), *Studies in Mon-Khmer and Thai phonology and phonetics in honor of E. Henderson*, 49–59. Bangkok: Chulalongkorn University Press.

Dingemanse, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6.10, 654–672.

Dingemanse, Mark. 2013. Ideophones and gesture in everyday speech. *Gesture* 13.2, 143–165.

Dingemanse, Mark, Will Schuerman, Eva Reinisch, Sylvia Tufvesson & Holger Mitterer. 2016. What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language* 92.2, e117–e133.

Dumestre, Gérard. 1998. Les idéophones: le cas du Bambara. *Faits de Langues: Revue de Linguistique* 11, 321–333.

Enfield, N. J. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.

Feinstein, Alvan R. & Domenic V. Cicchetti. 1990. High agreement but low Kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology* 43.6, 543–549.

Fortune, G. 1962. *Ideophones in Shona: An Inaugural Lecture given in the University College of Rhodesia and Nyasaland on 28 April 1961*. London & New York: Oxford University Press.

Fudge, Erik. 1970. Phonological structure and 'expressiveness'. *Journal of Linguistics* 6.2, 161–188.

Goldin-Meadow, Susan. What the hands can tell us about language emergence. *Psychonomic Bulletin & Review*, doi:10.3758/s13423-016-1074-x. Published online by Springer, 1 July 2016.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey* (Empirical Approaches to Language Typology 34). Berlin: Mouton de Gruyter.

Hamano, Shoko S. 1998. *The sound-symbolic system of Japanese*. Stanford, CA: CSLI Publications.

Havránek, Bohuslav. 1964. The functional differentiation of the standard language. In Paul L. Garvin (ed.), *A Prague School reader on esthetics, literary structure, and style*, 3–16. Washington, DC: Georgetown University Press.

Hinton, Leanne, Johanna Nichols & John J. Ohala (eds.). 1994. *Sound symbolism*. Cambridge: Cambridge University Press.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203.3, 89–96.

Joseph, Brian D. 1994. Modern Greek *ts*: Beyond sound symbolism. In Hinton et al. (eds.), 222–236.

Kageyama, Taro. 2007. Explorations in the conceptual semantics of mimetic verbs. In Bjarke Frellesvig, Masayoshi Shibatani & John Smith (eds.), *Current issues in the history and structure of Japanese*, 27–82. Tokyo: Kurosio Publishers.

Kakehi, Hisao. 1986. The function and expressiveness of Japanese onomatopes. *Bulletin of the Faculty of Letters, Kobe University* 13, 1–12.

Kelly, Spencer D. In press. Exploring the boundaries of gesture–speech integration during language comprehension. In R. B. Church, Martha W. Alibali & Spencer D. Kelly (eds.), *Why gesture? How the hands function in speaking, thinking and communicating*. Amsterdam: John Benjamins.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, Sotaro. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35, 379–415.

Klamer, Marian. 2002. Semantically motivated lexical patterns: A study of Dutch and Kambera expressives. *Language* 78.2, 258–286.

Kockelman, Paul. 2010. *Language, culture, and mind*. Cambridge: Cambridge University Press.

Kruspe, Nicole. 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press.

Kunene, Daniel P. 1965. The ideophone in Southern Sotho. *Journal of African Languages* 4, 19–39.

Kunene, Daniel P. 2001. Speaking the act: The ideophone as a linguistic rebel. In Voeltz & Kilian-Hatz (eds.), 183–191.

Le Guen, Olivier. 2012. Ideophones in Yucatec Maya. *Proceedings of the SSILA V Conference* 26, Austin, TX.

Levelt, Willem J. M. 1981. The speaker's linearization problem. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences* 295.1077, 305–315.

McNeill, David. 1992. *Hand and mind*. Chicago, IL: University of Chicago Press.

Mihas, Elena. 2012. Ideophones in Alto Perene (Arawak) from Eastern Peru. *Studies in Language* 36.2, 300–344.

Nasu, Akio. 2002. *Nihongo onomatope-no gokeisei-to inritu-koozo* [Word formation and prosodic structure of Japanese mimetics]. Ph.D. dissertation, University of Tsukuba.

Newman, Paul. 2001. Are ideophones really as weird and extra-systematic as linguists make them out to be? In Voeltz & Kilian-Hatz (eds.), 251–258.

Newmeyer, Frederick J. 1992. Iconicity and generative grammar. *Language* 68.4, 756–796.

Nuckolls, Janis B. 1996. *Sounds like life: Sound-symbolic grammar, performance, and cognition in Pastaza Quechua*. New York: Oxford University Press.

Nuckolls, Janis B. 2000. Spoken in the spirit of gesture: Translating sound symbolism in a Pastaza Quechua Narrative. In Joel Sherzer & Kay Sammons (eds.), *Translating native Latin American verbal art*, 233–251. Washington, DC: Smithsonian Press.

Nygaard, Lynne C., Debora S. Herold & Laura L. Namy. 2009. The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science* 33.1, 127–146.

Özyürek, Aslı. 2014. Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 369.1651, 20130296.

Peirce, Charles S. 1955. Logic as semiotic: The theory of signs. *Philosophical writings of Peirce*, 98–119. New York: Dover Publications.

Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33.2, 165–198.

R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Reiter, Sabine. 2012. *Ideophones in Awetí*. Ph.D. dissertation, Universität zu Köln.

Samarin, William J. 1970. Inventory and choice in expressive language. *Word* 26, 153–169.

Schultze-Berndt, Eva. 2001. Ideophone-like characteristics of uninflected predicates in Jaminjung (Australia). In Voeltz & Kilian-Hatz (eds.), 355–373.

Sidnell, Jack. 2006. Coordinating gesture, talk, and gaze in reenactments. *Research on Language & Social Interaction* 39.4, 377–409.

Slama-Cazacu, Tatiana. 1976. Nonverbal components in message sequence: "Mixed syntax". In William C. McCormack & Stephen A. Wurm (eds.), *Language and man: Anthropological issues*, 217–227. Berlin: Mouton.

Son, Youngsuk. 2010. Giongo/gitaigo-to miburi: Terebi-hoosoo-no maruchimedia-koopasu-ni yoru keiryoo-teki-bunseki [Relation of Japanese mimetics and gestures: A multimedia corpus-based quantitative study of Japanese television broadcasts]. *Keiryoo kokugogaku* [Mathematical linguistics] 27.4, 131–153.

Tamori, Ikuhiro. 1984. Japanese onomatopoeias: Manner adverbials vs. resultative adverbials. *Jimbun ronshu: Journal of Cultural Science* 20.2, 163–178.

Tamori, Ikuhiro. 1990. Expressiveness of Japanese and English onomatopoeic expressions. *Kotoba-no utage: Kakehi Hisao-kyoozyu kanreki kinen ronsyuu* ['Linguistic fiesta': Festschrift for Professor Hisao Kakehi's sixtieth birthday], 287–306. Tokyo: Kurosio Publishers.

Tamori, Ikuhiro & Lawrence Schourup. 1999. *Onomatope: Keitai-to imi* [Onomatopoeia: Form and meaning]. Tokyo: Kurosio Publishers.

Toratani, Kiyoko. 2006. On the optionality of *to*-marking on reduplicated mimetics in Japanese. In Timothy J. Vance & Kimberly Jones (eds.), *Japanese/Korean linguistics* 14, 415–422. Stanford, CA: CSLI Publications.

Toratani, Kiyoko. 2007. An RRG analysis of manner adverbial mimetics. *Language and Linguistics* 8.1, 311–342.

Toratani, Kiyoko. 2015. Iconicity in the syntax and lexical semantics of sound-symbolic words in Japanese. In Masako K. Hiraga, William J. Herlofsky, Kazuko Shinohara & Kimi Akita (eds.), *Iconicity in language and literature*, vol. 14, 125–141. Amsterdam: John Benjamins.

Tsujimura, Natsuko. 2005. A constructional approach to mimetic verbs. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical constructions: Back to the roots*, 137–154. Amsterdam: John Benjamins.

Tsujimura, Natsuko & Masanori Deguchi. 2007. Semantic integration of mimetics in Japanese. *The Main Session: Papers from the Thirty-Ninth Annual Meeting of the Chicago Linguistic Society* (CLS 39), 339–353. Chicago, IL: Chicago Linguistic Society.

Voeltz, F. K. Erhard & Christa Kilian-Hatz (eds.). 2001. *Ideophones* (Typological Studies in Language 44). Amsterdam: John Benjamins.

Watson, Richard L. 2001. A comparison of some Southeast Asian ideophones with some African ideophones. In Voeltz & Kilian-Hatz (eds.), 385–405.

Westermann, Diedrich Hermann. 1927. Laut, Ton und Sinn in westafrikanischen Sudansprachen. In Franz Boas (ed.), *Festschrift Meinhof*, 315–328. Hamburg: L. Friederichsen.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1987. Plain morphology and expressive morphology. In John Aske, Natasha Beery, Laura Michaelis & Hana Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* (BLS 30), 330–340. Berkeley, CA: Berkeley Linguistics Society.

Authors' addresses: *(Dingemanse)*
Language & Cognition Department,
Max Planck Institute for Psycholinguistics, P.O. Box 310,
6500 AH Nijmegen, The Netherlands
mark.dingemanse@mpi.nl

(Akita)
School of Languages and Cultures,
Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi,
Aichi 464-8601, Japan
akita@lang.nagoya-u.ac.jp