

Ideophones and gesture in everyday speech

Mark Dingemans

Max Planck Institute for Psycholinguistics, Nijmegen

This article examines the relation between ideophones and gestures in a corpus of everyday discourse in Siwu, a richly ideophonic language spoken in Ghana. The overall frequency of ideophone-gesture couplings in everyday speech is lower than previously suggested, but two findings shed new light on the relation between ideophones and gesture. First, discourse type makes a difference: ideophone-gesture couplings are more frequent in narrative contexts, a finding that explains earlier claims, which were based not on everyday language use but on elicited narratives. Second, there is a particularly strong coupling between ideophones and one type of gesture: iconic gestures. This coupling allows us to better understand iconicity in relation to the affordances of meaning and modality. Ultimately, the connection between ideophones and iconic gestures is explained by reference to the depictive nature of both. Ideophone and iconic gesture are two aspects of the process of depiction.

Keywords: ideophones, gesture, depiction, iconicity, iconic gesture, composite utterances

The ideophone is the closest linguistic substitute for a non-verbal, physical act.
Daniel P. Kunene (2001)

Introduction

Many of the world's languages have large numbers of ideophones. These are words that are marked relative to ordinary words, characterised by peculiar phonotactics (e.g. skewed phoneme distributions, feature harmony) and expressive morphology (reduplication and lengthening). Speakers tend to produce and interpret them as performances: words that do not simply refer by description but that “act out” or “depict” meaning using various types of iconicity (Dingemans, 2011a; Goodman, 1968; Güldemann, 2008). Their meanings are in the domain of sensory imagery, evoking all sorts of perceptions and inner sensations. Accordingly, ideophones are defined as *marked words that depict sensory imagery*. To give some examples, we

have, in Japanese, *gorogorogoro* ‘rolling’, *uja-uja* ‘insects swarming’ and *pikapika* ‘shiny’ (Kita, 1997) or in Siwu *sinisinisini* ‘closely woven’, *mukumuku* ‘mumbling mouth movements’ and *saaa* ‘cool sensation’ (Dingemans, 2011b). Across languages, ideophones tend to be relatively syntactically independent, delivered with prosodic foregrounding that sets them off from the rest of the utterance. They are often introduced by quotative markers or “say” or “do” verbs, emphasizing their performative nature. All of these features work together to mark ideophones as depictions, much as the frame around a painting tells us to interpret it as a painting and not as the wallpaper.

A special relation between ideophones and gestures has often been reported. For example, Kunene noted that in Sesotho, a Southern African Bantu language “[m]any ideophones are often accompanied by gestures imitating the action or state conveyed by the ideophone” (Kunene, 1965, p.32); Zondo noted that in Ndebele, a Bantu language spoken in Zimbabwe gesture “almost always accompanies an ideophone” (Zondo, 1982, p. 123). While these claims were limited to particular languages, Kita, examining data from Japanese and Mandarin, elevated the tight coupling of ideophones (or ‘mimetics’ as he calls them) and gesture to universal status: “constant and fairly accurate synchronisation between mimetics and gestures is universal among the languages that have mimetics” (Kita, 1993, p. 96).¹

This paper examines these proposals with two goals: (1) to see how they hold up in a corpus of everyday conversation; (2) to provide a more explicit account of *why* there might be a special relationship between ideophones and gesture. Along the way, we will also see what the co-occurrence of ideophones and iconic gestures can teach us about iconicity. The empirical data is a corpus of naturally occurring conversations in Siwu, a Kwa language spoken in Ghana, West-Africa, in the region where ideophone systems were first described.

Previous work on ideophones and gesture

The earliest claims about the link between ideophones and gesture date from the pre-video era. Though there was general agreement about some kind of connection between ideophones and gesture, few actual examples were available. Schlegel (1857, p. 114), in a study of Ewe, a Kwa language of present-day Ghana, saw ideophones and gestures as both arising from “an urge to illustrate all conceivable shades of meaning”. Hetherwick (1889, p. 76), in a grammar of the Yao language of present-day Tanzania, described an ideophone *myu* signifying ‘completion’, and produced “with the fingers drawn across the lips, or accompanied by a peculiar motion of the hands one over the other”. This description was cited in Werner (1919, p. 187), and Werner’s statement was in turn cited by Samarin in support of

the possibility that “certain ideophones are always accompanied by, for example, a certain gesture” (Samarin, 1971, p. 153).

More detail was offered in a seminal publication on ideophones in Southern Sotho by Daniel Kunene, native speaker of the language. Kunene (1965) was one of the first to argue for ideophones as fundamentally performative words. He saw them as performances in which speakers “act out” the meaning, conjuring up sensory events for listeners. This provided a natural fit between ideophones and gestures, especially gestures of the representational or iconic type. As he wrote, “Many ideophones are often accompanied by gestures imitating the action or state conveyed by the ideophone, more so in the case of actions than of states. Sometimes, where possible, the speaker actually turns actor and performs the action referred to” (Kunene, 1965, p. 36). Kunene’s work, like that of his contemporaries Fortune (1962) and Zondo (1982), was strongly influenced by the work of Doke on southern Bantu languages (Doke, 1935), which saw ideophones as a rhetorical device most at home in narratives. We will see later how discourse type is relevant to the study of ideophones and gesture.

In the second half of the 20th century the study of ideophones took a more cross-linguistic turn. Diffloth, working on ideophones in South-East Asian languages, noted that when speakers are asked to explain ideophones, “many speakers cannot find exact paraphrases and prefer to repeat the ideophone with a more distinct elocution, accompanied by facial expressions and body gestures if appropriate. This sort of behavior does not occur when we ask the same persons to paraphrase Nouns or Verbs — a first indication that ideophones have a mode of meaning which is different from that of ordinary words” (Diffloth, 1972, p. 441). Samarin offered similar observations from his work on ideophones in Gbaya, a Niger-Congo language of central Africa: “It turned out that some of the meanings I isolated were based almost exclusively on gestures. On the assumption that informants were leaning too heavily on their gestures to convey the meanings, I have tried, unsuccessfully to get them to verbalize without gestures. More sophisticated experiments should be done; gestures, of course, are not to be outlawed entirely from investigations!” (Samarin, 1971, p. 154).

Janis Nuckolls, in a productive line of work on ideophones in Pastaza Quechua, has written about the gestures that co-occur with ideophones in narratives during interviews, as well as about the sense in which ideophones themselves may be said to be vocal gestures (Nuckolls, 2000, 1996). Reiter, in an unpublished PhD thesis on Awetí, a Tupian language spoken in Brazil, noted that ideophones “seem to be nearly always accompanied by gestures” (Reiter, 2012, p. 404), and found that ideophones co-occur especially commonly with pointing and iconic gestures in a varied corpus of myths, task descriptions, and explanations. The descriptive work by the authors cited has not always distinguished what speakers do in explanations

or interviews from what speakers do in everyday conversations. It is important to keep this distinction in mind, for the simple reason that when explaining an ideophone, speakers may be doing something quite different than when using that same ideophone in everyday speech.

A number of recent studies have focused on the relation between ideophones and gesture in rather specific contexts of use. Klassen (1999) provided a beautifully detailed single case analysis of ideophones and gesture in a performance of the *ngano* story-song genre in Shona. She argued for linking the acoustic and kinesthetic dimensions of ideophones to performance-generated meaning in a way that is reminiscent of Kunene's identification of ideophones with performances, but with much stronger empirical grounding. Kita (1997, 1993) studied Japanese ideophones and gestures in a corpus of retellings of Tweety cartoons. He found a particularly tight coupling of ideophones and iconic gestures: 94% of ideophones (N = 83) co-occurred with an iconic gesture (Kita, 1993, p.90). Taking up an argument first introduced by Diffloth, Kita used this finding as evidence for the idea that ideophones employ a mode of meaning that is different from ordinary words. He called this the "affecto-imagistic" mode of meaning. Seen against the background of the turn in cognitive science toward incorporating bodily experience and action as a foundation for cognitive processes, Diffloth's and Kita's arguments can be re-interpreted as pointing to the embodied, sensory meanings of ideophones.

In sum, then, pre-video era studies offer acute, but anecdotal observations; later descriptive studies have contributed to our understanding of ideophones as depictive-performative words, but focused mostly on gestures in explanations of ideophones; and the most detailed empirical studies to date have investigated the gesture-ideophone relation in rather narrow genres of elicited narratives. One thing that is missing is a study of the relation between ideophones and gesture in everyday face to face conversations — the primordial home of language (Levinson, 1983; Schegloff, 2006), and the baseline against which other, more specialised genres are best compared (Dingemans, 2011a, ch. 11–12). The present study is designed to fill that void. Its primary aim is rather modest: it is to see whether some of specific, commonly repeated claims about ideophone and gesture hold up in a video corpus of everyday speech — claims like "ideophones are often accompanied accompanied by gestures of mimicry" (Kunene, 1965, p.21), "[the gesture] almost always accompanies an ideophone" (Zondo, 1982, p.123), and "mimetics [ideophones, MD] are tightly coupled with spontaneous iconic gestures" (Kita, 1997, p.392). There are commonly two aspects to these claims, which may be usefully posed as research questions:

- I. *Gesture type*: is there a special relationship between ideophones and a certain type of gesture? If so, what is its nature and cause?

II. *Frequency*: are “often” and “almost always” fair characterisations of ideophone-gesture couplings in everyday conversations?

Methods and data

A video corpus of informal, naturally occurring Siwu conversations was collected in Ghana according to standard methods in the field study of social interaction (Enfield, 2013). A subset of about 3000 utterances was transcribed and analysed in detail (Dingemanse, 2011a). The unit of analysis for this particular study is the communicative move, the basic-level unit for social interaction (Enfield, 2009, pp. 11–12; Goffman, 1981). The type of communicative move in focus is a *composite utterance* (Clark, 1996; Enfield, 2009) in which speech and gesture are combined, and specifically that kind of composite utterance in which we can study the relation between ideophone and gesture. The corpus contained 219 utterances featuring ideophones. Of these, 174 are uttered by a speaker visible in the frame, and these form the corpus for this study.

The 174 composite utterances in the corpus were coded for gestural activity co-occurring with ideophones. To be considered, the stroke (meaningful phase) of a gesture had to coincide with the ideophone. Cases in which ideophone and gesture stroke only coincided for a brief amount of time made up no more than 5% of the total, and were not considered. Coding distinguished between four gesture types: pointing gestures, iconic gestures (also known as representational or depictive gestures), emblematic gestures, and beats (see McNeill, 1992; Kendon, 2004, pp. 84–107 for discussions of these gesture types). Despite the intrinsic interest of kinesic details, further possible subclassifications of gesture, and precise time measurements of gestures and the words associated with them (or words and the gestures associated with them), this study will not address these matters in order to keep a view of the primary goal: to assess a number of commonly repeated, wide-ranging claims about ideophones and gesture, specifically with regard to frequency and gesture type.

Results

In the corpus, all four types of gestures were found co-occurring with ideophones, though with different frequencies. Before discussing the coding results it will be useful to discuss examples of the four gesture types as they occur in ideophonic utterances.

Ideophones and gesture types

Iconic gestures depict aspects of the same scene that the speech also represents. An example from the corpus is (1) below, in which the speaker explains what is going to happen when he sets fire to two small amounts of drying gunpowder: they will burst into flames *shû*, *shû*.

- (1) ò-bra ì-a-bra shû shû
 PFOC-make it-FUT-make IDPH.flare.up IDPH.flare.up
 | G1 | | G1 |

“It’ll go *shû*, *shû*”

G1: both hands moving symmetrically in a quick upward motion; RH holds an object, LH is spread out palm upward

The form of the ideophone *shû* depicts several aspects of the flaring event: the fricative onset mimics the sizzling sound of the powder flaring up, the brief monosyllabic form mimics the brief duration, and the falling intonation can be seen as an iconic diagram of the intensity of the burst. The speaker also depicts the flaring up of the flames gesturally, by moving both hands symmetrically in a quick upward motion. Ideophone and iconic gesture are produced in tight synchrony, twice in quick succession.

Pointing gestures are gestures that indicate either concrete or imaginary objects or people. An example is (2), in which the speaker tells the story of a king who was taking baths with hot and cold water. Two gestures indicate the location of the imaginary buckets with hot and cold water, the first pointing up and to the speaker’s right and the second pointing up and to the speaker’s left. The first gesture is aligned with the noun phrase *ndu tɔtɔ-à* ‘hot water’, the second with the nominalised deideophonic adjective *miyululu-à* ‘cold one’.

- (2) n-du tɔtɔ-à gu mi-yululu-à
 MI-water hot-ADJ with AGR.MI-IDPH.cold-ADJ
 | G1 | | G2 |

“hot water and cold [water]”

G1: RH pointing outwards from the speaker’s upper right periphery

G2: RH pointing outwards from the speaker’s upper left periphery

Emblems (following a definition widely adopted from Ekman & Friesen, 1969) are highly conventionalised gestures that can be said to have a lexicalised meaning. A stock example from English is the gesture joining index finger and thumb, with the other fingers outstretched, which means “excellent/OK”. An example from Siwu is a quick waving of the hand or hands in front of the body, palm outward, for “no” or “not at all”. In (3), the speaker concludes a narrative of personal hardship with the assurance that “it wasn’t easy!” while producing this emblem. The stroke of the

gesture extends over the whole of the utterance, including the ideophone, indicating the scope of negation.

- (3) *ì-i-se* *pəkəsəwə!*
 I-NEG-be IDPH.slow/easy
 | G1 |
 “it wasn’t easy at all!”

G1: repeated waving of the right hand, palm outward, in front of the body

Finally, *beats* are simple rhythmic hand movements which appear to bear a relation to prosodic features or discourse-level meanings more than to the meanings of lexical items. In example 4 below, two beats occur in the form of minimal wrist movements in which the hand (palm upward) goes up and down. An upward movement starts somewhat before *sí* with a high point at *ló*; a downward movement brings the hand to resting position again at *fufu*. Then again an upward movement up to and during *qəbərəwə* and a downward movement right after, with the hand back at resting position at *ə-se*. The selected utterance is part of a longer stretch during which this simple biphasic beat pattern is repeated several times.

- (4) *sí ló-wètè* *fufu, qəbərəwə* *ə-bià* *ə-se* *ə-qə...*
 if 1SG:PST-pound *fufu*, IDPH.soft 3SG-be.very 3SG-HAB 3SG-love
 | G1 | | G1 |
 “when I pounded *fufu*, soft as he very much liked it...”

G1: peak of slight wrist movement moving the RH (palm upward) up

What these four examples show is that not all instances in which gesture and ideophone co-occur are alike. The iconic gestures in (1) are produced in synchrony with the ideophone *shú* ‘flaring up’, and they appear to depict in the visuo-spatial modality aspects of what the ideophone depicts in the auditory modality. Here the ideophone could be said to be the lexical affiliate (Schegloff, 1984) of the gesture (or vice versa, the gesture might be said to be the gestural affiliate of the ideophone): there is a solid one-to-one correspondence between the two, as seen from the temporal alignment of gesture stroke with ideophone, the precise repetition of both, and the fact that both depict aspects of one and the same scene. The pointing gestures in (2) point out the locations of states of affairs expressed by two noun phrases, one of which is derived from an ideophone. The temporal co-occurrence is again straightforward, but the link between gesture and speech is now a deictic one: the pointing gestures set up an imaginary space in which the listener is to locate the two buckets of water, one hot and one cold.

In the last two examples there is a less tight relation both in temporal and semiotic terms. The emblematic gesture in (3) fully overlaps with the ideophone, but its placement is more properly expressed in terms of its relation to the utterance as

a whole: the gesture expresses negation scoped over the whole expression and as such has little to do with the ideophone as such. Finally, the beats in (4) bear only the most tenuous relation to the ideophone. While one of the beats in (4) is synchronised with an ideophone, this may have to do more with the prosodic rhythm of the utterance than with the ideophonic nature of *qoborɔɔ*. For the purposes of this study, the cases that combine both temporal co-occurrence and a privileged semiotic link will be the most interesting, as they will shed light on the ideophone-gesture relation most clearly.

Quantitative results

With the examples above in mind, we can now take a quantitative look at the data. Figure 1 shows the frequency of the four gesture types co-occurring with ideophones in composite utterances.

Although all gesture types are found in temporal co-occurrence with ideophones, by far the strongest relation is between ideophones and iconic gestures: of the 66 gestures identified as co-occurring with ideophones, 52 (or 79%) are iconic, while the other three types together amount to only 14 (or 21%). In other words, *if* an ideophone comes together with a gesture in everyday speech, it is highly likely to be an iconic gesture. This answers the first of our research questions, that of gesture type: there is indeed a strong relationship between ideophones and iconic (representational, depictive) gestures.

What about frequency? If prior claims about the frequency are indicative of the situation in naturally occurring speech, we should find that ideophones are almost always accompanied by gestures. Recall for instance that Kita (1993) showed that 94% of ideophones in a Japanese corpus of cartoon retellings came with a gesture. Table 1 shows a count of the ideophones with and without gesture in the

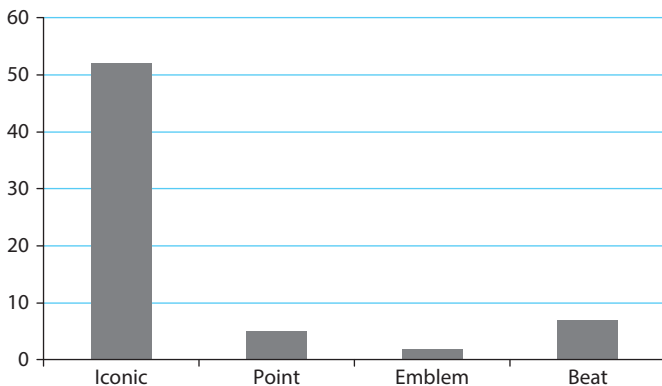


Figure 1.

Table 1. Ideophones with and without gesture in the Siwu corpus

| <i>Ideophone with...</i> | Gesture | No gesture | Total |
|--------------------------|----------------|-------------------|--------------|
| Tokens | 66 | 108 | 174 |

Siwu conversational corpus: about 38% of all ideophones come with a gesture. While this is a substantial chunk of cases, it does not nearly come close to the 94% found by Kita (1993) or to the “almost always” of Zondo (1982). This appears to be a major blow to strongly phrased claims about the near-universal co-occurrence of ideophones and gesture.

However, perhaps gesture-ideophone co-occurrences are more common in some types of talk than in others. In fact, since prior work has mostly been based on elicited narratives, a more specific hypothesis can be formulated: gesture-ideophone co-occurrences are especially common in narrative contexts. To test this hypothesis, the data was further coded to distinguish between tellings and other activity types. Tellings are here defined as multi-unit turns in which a speaker produces a temporally ordered report of some experience or event (Goodwin & Heritage, 1990, pp. 298–300 and references therein).

Table 2 breaks down the gesture-ideophone co-occurrences according to whether they occur in a stretch of discourse that can be qualified as a telling or not. The amount of gesture-ideophone couplings is much higher in tellings than in other contexts: 53% as opposed to 30%. The difference in gesture-ideophone co-occurrences between these two contexts is statistically significant ($p = 0.0052$, Fisher’s exact test).

Table 2. Ideophones with and without gesture in tellings vs. other discourse contexts

| <i>Ideophone with...</i> | Gesture | None | Total |
|--------------------------|----------------|-------------|--------------|
| Telling | 31 | 28 | 59 |
| Other | 35 | 80 | 115 |

While 53% is still far from “almost always”, the difference between tellings and other contexts with regard to ideophone-gesture co-occurrence points the way to an explanation of earlier claims. These claims were likely based on data heavily skewed towards narratives and explanations, both contexts in which more gestures may be expected.

Discussion

We have seen that *if* an ideophone comes together with a gesture, the gesture is likely to be an iconic gesture; and that the link between ideophones and gesture is especially strong in tellings. Starting with the latter finding, why would discourse type make a difference, and why would stories be inductive to the use of ideophones together with iconic gestures?

One reason may be that narratives have as one of their main goals letting the hearer know what it was like to experience the events reported (Cassell & McNeill, 1991; Stivers, 2008). Such a goal invites the liberal use of depictive devices like ideophones and iconic gestures. This may explain not just the strong association of ideophones and iconic gestures in narratives, but also the often made observation, not directly addressed in this paper, that ideophones appear to be relatively more common in narrative contexts. In many non-narrative contexts in contrast, the main business may be something else. Speakers may be carrying out a conversational routine that does not call for much performative elaboration; or even simpler, they may be engaged in a joint activity in which their hands may not be free. Such factors would decrease the number of ideophone-gesture co-occurrences, explaining the difference we find in tellings versus other discourse contexts.

With regard to ideophone-gesture couplings within narrative data, there is still the difference between how commonly ideophones come together with gesture: 54% in the Siwu corpus and 94% in the Japanese corpus (Kita, 1997, 1993). This difference may in large part be due to further qualitative differences between the corpora. After all, spontaneous tellings in conversation are potentially different along several parameters from elicited retellings of cartoon films. For instance, it has been noted that the lively Sylvester-Tweety Bird cartoon underlying Kita's narrative corpus may be especially prone to elicit certain ideophones and iconic gestures (Tsujimura, 2001).² Kita (2001) has acknowledged this possibility, noting that the ideophones in his corpus of cartoon retellings were almost exclusively in the spatial domain. The Siwu ideophones in the present corpus are much more varied in meaning, and some of them may be less easily accompanied by iconic gesture than others, a possibility explored in the section on iconicity below.

Why ideophones and iconic gestures go well together

While it may be fair to conclude that there is a good deal of variation in the overall frequency of ideophone-gesture co-occurrences, a substantial finding confirmed here is nonetheless that there *is* in fact a robust and particularly tight coupling between ideophones and iconic gestures. Why would this be?

Two suggestions stand out in prior work. The first is Kunene's observation that ideophones are essentially performances. As he noted, "The most remarkable thing about the ideophone is that it constitutes a dramatic presentation of past and future events and states. By means of it, the speaker conjures up before the eyes of his listeners events which are not actually happening, or states which are not existing, at the time of speaking. It is for this reason that ideophones are so often accompanied by gesture of mimicry and sometimes indeed completely replaced by them" (Kunene, 1965, p. 21). The second suggestion is a more cognitively oriented explanation by Kita, according to which the coupling of ideophone and gesture is due to the "inherent imagistic meaning" of ideophones. Kita, a native speaker of Japanese, also brings up "native speakers' intuition that mimetics evoke imagery" (2001, see also 1997, 1993, p. 98). These two perspectives — Kunene's outward-looking, behaviour-linked one and Kita's inward-looking, cognition-linked one — are best thought of as two sides of the same coin. The kernels of insight they embody can be reframed in terms of two reasons for the common coupling of ideophones and iconic gesture: both ideophone and iconic gesture form part of a single performative act; and both ideophone and iconic gesture are holistic depictions of complex states of affairs. Let me discuss both in turn.

The coupling of ideophone and iconic gesture may be thought of as the temporal co-occurrence of depictive signs across modalities. Recall that ideophones have a fundamentally depictive mode of representation, something they share with iconic gestures (Kendon, 2008; Streeck, 2008). Now, ideophones are part of the linear, one-way flow of the unfolding speech signal. The dominant mode of signification for lexico-syntactic speech is descriptive rather than depictive. To exploit verbal material for depictive purposes it is therefore necessary to signal a shift in mode of signification — from descriptive to depictive. As argued in more detail elsewhere (Dingemanse, 2012), the position of ideophones (at clause edges, often set off with an intonational break), their structural markedness, their expressive morphology and their prosodic foregrounding all conspire to serve this goal. Thus there is usually a clear point at which, in Kunene's insightful turn of phrase, the speaker turns actor. From this point of view, the coupling of ideophone and iconic gesture is entirely natural: it is not a case of independent units in disparate modalities being synchronised, but of a speaker-turned-actor using all available means to produce a single, richly multimodal act of depiction.

Support for this suggestion comes from the fact that the coupling of ideophone and iconic gesture is most common when the ideophone is found at clause edge, showing prosodic foregrounding and expressive morphology, and that it is weak or even absent for ideophones that are deeply embedded in utterances and burdened by descriptive morphology (Dingemanse, in press). Examples 5 and 6 illustrate. Both feature the ideophone *dɔbɔrɔ* 'soft'. In the first, the ideophone occurs

at clause edge, a position where it enjoys ample freedom: it is prosodically foregrounded and comes with some extra lengthening (a case of expressive morphology). It also comes together with an iconic gesture depicting the soft, fine-grained nature of the gunpowder being described. In the second, the same ideophone is more integrated into the morphosyntax: it occurs in a backgrounded, reasons-giving clause as a predicate with preverbal morphology. When morphosyntactically integrated like this, it does not undergo expressive morphology, and does not come with an iconic gesture. The ideophone is now used as an ordinary verb, and being no longer a depictive act, it is not accompanied by an iconic gesture.

- (5) *kōrō ne, kùwà gɔ-ŋgbe kù-nyɔ ↑dɔbɔrɔwɔwɔ↑*
 now TP C.KU-stuff A.KU-DPRX S.KU-look IDPH.SOFT.EM
 |G1 | | G2 |
 “Now this stuff here, it looks *dɔbɔrɔwɔwɔ* [soft, fine-grained]”
 G1: RH points to mortar in which gunpowder is being fabricated
 G2: RH moves down slowly, fingers spreading out and then coming together as if scooping up powder
- (6) *igɔ̃ ne bo gu Tasi ɔ-kā ne kà-ɔ-dɔbɔrɔwɔ wso ne*
 TDBY TP 1PL with PSN SCR:PST-squeeze TP ING-SCR-IDPH.SOFT reason TP
 “The day before yesterday, me and Tasi wrung [the palm fruit pulp]³ because it was getting soft.”

The second way of thinking about the coincidence of ideophones and iconic gestures — the other side of the coin — is that ideophones are likely lexical affiliates for iconic gestures. They are in this respect unlike some other types of words. It has been observed by several authors that it is often difficult to point to a single lexical affiliate for iconic gestures. For instance, De Ruiter has argued that “many iconic gestures reveal properties that can, at best, only be represented by phrases” (de Ruiter, 2000, p. 291). He discusses a pantomimic gesture from the McNeill lab cartoon retellings that “reveals aspects of the bowling ball itself, of holding it, of throwing it, and of throwing it in a downwards direction” (*ibid.*), while in the accompanying speech there is no single lexical affiliate for this complex depiction. Accordingly, De Ruiter suggests that iconic gestures do not have lexical affiliates but rather affiliate with conceptual representations. “All content words have an underlying conceptual representation, but not all conceptual representations have a corresponding content word” (*ibid.*). This would explain that there is often no single lexical affiliate for an iconic gesture, but it does allow for the possibility that sometimes words and iconic gestures temporally coincide.

Here the special nature of ideophones becomes relevant. If, as de Ruiter (2000, p. 291) says, iconic gestures “reveal properties that can, at best, only be represented by phrases”, in ideophones they meet their verbal counterpart. Ideophones too

have often been described as evoking scenes that can only be represented with phrases. For example, Diffloth compares ideophones to “microscopic sentences” (Diffloth, 1972, p. 444); for Noss, each ideophone “represents a full clause” (Noss, 2001, p. 267); and Güldemann notes that ideophones “evoke a whole state of affairs” (Güldemann, 2008, p. 280). Both iconic gesture and ideophone are simplex units depicting complex states of affairs. This is what makes ideophones likely lexical affiliates for iconic gestures. However, a better way to put it — better because it avoids expressing the relationship as one of static “affiliation” — is that both ideophone and iconic gesture point to and arise from a third thing: the the sensory imagery depicted by the multi-modal act.

A conjecture about ideophones and iconic gestures

With the above proposals about the relation between ideophones and iconic gesture in mind, we can return to a conjecture by Kita (1993) (based on evidence from Japanese and Mandarin Chinese) that the tight coupling between ideophones and iconic gesture may well be universal. While the present study provides some support for this suggestion, it also provides reason to modify and sharpen it. Recall that Kita phrased his conjecture in terms of a universal implicational relation between the existence of ideophones in a language and their synchronisation with gesture strokes:

(7) *Kita’s conjecture*

Constant and fairly accurate synchronisation between mimetics and gestures is universal among the languages that have mimetics. (Kita, 1993, p. 96)

Based on the evidence and arguments presented here, I propose to reframe and sharpen this conjecture. There are two main problems. First, it can be read as a blanket claim about all types of gestures, although from his discussion it is clear that Kita did not intend it as such. Read in this strong way the conjecture is simply false. The varied movements that have been called beats, for example, are not exclusively synchronised with any one word; if anything, their attachment appears to be with higher level discourse units. Emblems and pointing gestures may sometimes co-occur with ideophones, but there is no semiotic or semantic reason why there would be a privileged relationship between these gesture types and ideophones. The data presented above show that the really special relation is between ideophones and iconic gestures.

Another drawback is that the conjecture is framed in terms that obscure the nature of the processes at hand. For instance, the term “synchronisation” invites the interpretation that ideophone and iconic gesture are independent entities in disparate modalities that happen to become temporally aligned with each other

due to some unspecified force. But as argued above, the temporal relationship between ideophone and iconic gesture is not a mere statistical correlation; its ultimate cause is well understood and can be precisely specified. The temporal coincidence of ideophone and iconic gesture is due to the fact that they are two aspects of one multi-modal performative act of depiction. We thus arrive at the following revised conjecture:

(8) *The relation between ideophones and iconic gestures*

Ideophone and iconic gesture are two aspects of a depictive performance, and as such will tend to coincide temporally within an utterance.

This conjecture captures the semiotic and temporal relationships between ideophone and iconic gesture. It reflects the fact that the temporal co-occurrence of ideophone and iconic gesture is a consequence of a speaker-turned-actor mobilising all expressive resources in the act of depicting some sensory imagery. Phrased this way, the conjecture seems defensible as well as falsifiable. In this article, some headway has been made towards supporting it. However, the matter is not entirely settled. There are well-known problems in deciding what exactly it means for speech and gesture to be temporally coinciding. Summarizing from de Ruiter (2000, pp. 297–300), the problems include first of all the fact that it can be difficult to decide on the lexical affiliate of a gesture; second, the question of how to define temporal coincidence (Butterworth & Hadar (1989) enumerate 13 different ways in which two time intervals can be related); and third, practical difficulties in measuring the time intervals of gestures. Only the first of those is mitigated in the case of ideophones because, as I have argued, ideophone and gesture often clearly depict aspects of the same thing. Further research would be needed to tackle the other two in a satisfying way.

Note that unlike some previous claims, the conjecture is not directional. That is, it does not stipulate that iconic gestures come with ideophones or vice versa; it conjectures only that when the two are both present in an utterance, they will tend to occur together. It is worth briefly elaborating this. It appears that some authors have conceived of the relation between ideophone and gesture as asymmetric; thus when Zondo writes that the gesture “almost always accompanies an ideophone” and is “of vital importance to the overall semantic import of an ideophone” (1982, p. 123), it appears that he takes the gesture to be crucial for interpreting the ideophone, but not vice versa. This overlooks the fact that the interpretation of the gesture in the context of the utterance is also enriched by the ideophone.⁴ Thus, the process of comprehending words-in-context does not stop with mere recognition of lexical items; it also involves contextualized interpretation that takes into account the multiple signs of multiple types (including gestures) that make up the composite utterance. Likewise the process of comprehending gesture-in-context

takes into account the other types of signs (including ideophones) that are part of that utterance. In semiotic terms, both ideophone and iconic gesture point to a third thing; both are enriched by their co-occurrence, and together they are taken as signs of a single informative — or more precisely, performative — intention.

Iconicity

Both ideophones and iconic gestures are holistic, iconic depictions of sensory imagery. How do the iconic techniques they employ compare? Even though there is quite some work on depiction in speech (Westermann, 1927; Nuckolls, 1999; Dingemanse, 2011c) as well as on depiction in gesture (Mandel, 1977; Kendon, 2004; Streeck, 2008), there are few occasions where the two have been considered together (but see Perniss et al., 2010; Kendon, 2008). Iconicity — defined here as a perceived resemblance between form and meaning — is a semiotic relationship that comes in different kinds and depends on the affordances of medium and modality. When investigating how iconicity works in a certain kind of medium or for depicting a certain kind of meaning, it is useful to consider what we may call *the iconicity question* (revised from Dingemanse, 2011c):

(9) *The iconicity question*

What are the structural properties of form and meaning such that they afford iconic mappings between the two?

Iconicity relies on perceived resemblances. The structural properties of form and meaning, or sign and signified, play enabling as well as constraining roles in the construction of these resemblances. Gesture may afford the iconic expression of certain aspects of meaning (for instance, motion paths or shape), while speech may afford the iconic expression of other aspects (for instance, sound or intensity). But speech and gesture also partly overlap in their affordances for iconicity: both have internal structure in the form of duration and temporal unfolding, and both may thus be used for the iconic expression of aspects of meaning that share in this kind of structure (for instance, contrasts in aspectual semantics or duration).

The iconicity question, agnostic with regard to meaning and modality, captures multiple possible perspectives on iconicity. Sometimes it is useful to focus on how different modalities offer different affordances for iconicity, for instance when answering the question why sign languages seem to have more iconic lexical signs than spoken languages (Mandel, 1977) or when investigating the relative semantic contributions of iconic gestures versus linguistic material in conversation (Holler & Beattie, 2003). Sometimes it is useful to focus on how different meanings afford expression across modalities, for instance when looking at the diagramming of

resemblance to the flaring up of the flame. During the rapid motion, the turning of the hand from pointing downwards into a spread-out palm upward position can be seen to resemble the way in which the flames spread from around the point of ignition, with the hand becoming a model of the flame. The distance covered by the motion appears to represent the expected height of the flames, in accordance with the cautionary nature of the speaker's message. Thus, the gesture depicts the same flaring up event as the ideophone, but evokes different aspects of it in accordance with the different possibilities offered by a visuo-spatial medium.

If affordances can be a source of asymmetry, it follows that we may expect to find qualitative and quantitative differences in the nature of the iconic gestures accompanying ideophones, depending in part on the sensory imagery evoked. As it happens, Kita's (1993, 1997, 2001) study of Japanese ideophones in cartoon retellings can be construed as pointing to just that. His corpus — elicited retellings of an animated Sylvester and Tweety Bird cartoon — is essentially one that controls for ideophone type: the ideophones in it are “strongly skewed to spatial events and states” and there are “relatively few mimetics for nonspatial domains” (Kita, 2001, p. 425). In his corpus, as many as 94% of ideophones come with an iconic gesture. This is not surprising given the fact that gesture offers many affordances for depicting meanings in the domains of space and movement. The ideophones in the Siwu conversational corpus are much more varied, and while the present study was not designed to specifically address this matter, it does seem to be the case that ideophones in the domains of movement, size, and shape come with iconic gestures relatively more often than ideophones in other domains (e.g. sound, temperature, inner feelings). This is in line with the idea that the affordances of meaning and modality play a role in the recruitment of expressive resources.

Conventionalisation

The common coupling of ideophone and iconic gesture brings up the question of conventionalisation. Most ideophones are lexical items that have conventionalised meanings within a speech community; could the same hold for the iconic gestures that accompany them? Certainly some of the more frequent ideophones in Siwu come with gestures that are quite similar across instances of use. For instance the ideophone *gbògbòrò* ‘tough, strong’ often comes with a gesture of flexing the arm muscles, and the ideophone *minimini* ‘spherical’ often comes with gestures depicting round objects.

When the same ideophone is accompanied with similar gestures across instances, there are two main possibilities: (I) the gestures may look the same because they are attempts to depict the same kind of sensory imagery (this would be regularity stemming from *conceptual commonality*); or (II) they may look the

same because they are subject to some kind of socially mediated convention (this would be regularity regimented by *communicative convention*). The possibilities are not mutually exclusive, and it is of course easy to see that repeated regularity may over time turn into communicative convention by means of known processes of conventionalisation (Clark, 1996; Kendon, 1997).

It can be hard to investigate the conventionalisation of iconic gestures in a corpus of everyday speech because there is no guarantee that the same ideophone will be recorded enough times to permit close study of the accompanying gestures. A separately collected corpus of folk definitions of ideophones in Siwu does throw some light on the question. For instance, in that corpus, there are gestural depictions of the ideophone *minimini* 'spherical' by four different speakers (Figure 3). Despite important similarities — all of the gestures clearly depict some spherical object — there is a great deal of variation: while two speakers enact the sculpting of a medium-sized sphere, a third spreads out his arms relatively wide to model a large sphere and a fourth brings her cupped hands together four times to model four round types of fruit (Dingemans, 2011a, pp. 220–224). So the gestures differ at least in terms of size, position of the hands, and method of representation (Kendon's (2004, p. 160) "modelling" versus "enactment"). Given the amount of variety across these different parameters, it is likely that the similarities we see in these particular gestures are due to the commonality of the sensory imagery depicted rather than due to convention.

In sum, whereas the similarity of the iconic gestures accompanying ideophones sometimes may simply be due to a commonality in the imagery depicted, some of them may attain, by sheer force of habit, a degree of conventionalisation.

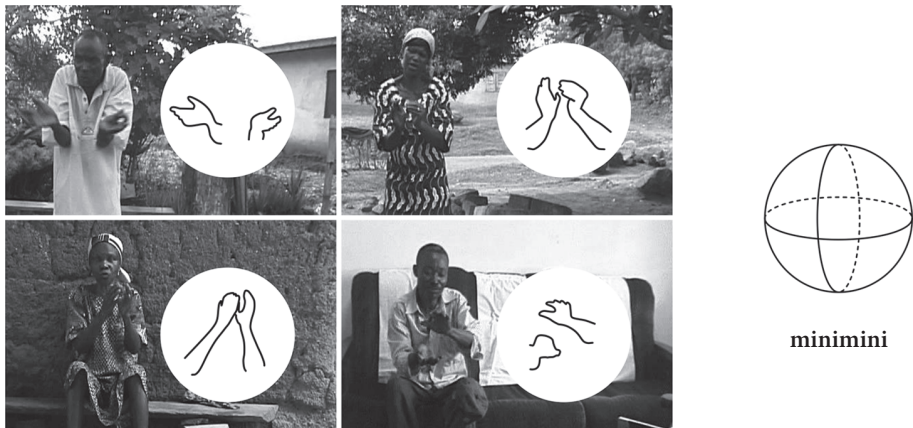


Figure 3.

Conclusions

This study has addressed two aspects of prior claims about the special relationship between ideophones and gesture. With regard to *gesture type*, we saw that there is indeed a particularly tight coupling between ideophones and iconic gestures. With regard to *frequency*, we saw that previous claims turn out to be too strong: it is not the case that gesture “almost always accompanies an ideophone” (Zondo, 1982, p. 123), certainly not in everyday language use. Discourse type makes a difference here: ideophones are more commonly accompanied with gestures in tellings than in other contexts, which may well explain the previous claims, since we know that the database of ideophone research has been heavily biased towards narratives. A general lesson therefore is that the common practice in ideophone and gesture research to work with elicited narratives rather than natural conversational data may lead to a skewed view of the actual (qualitative and quantitative) nature of the phenomena under study.

Despite the lower-than-previously-claimed frequency, the empirical findings of this study show that there *is* a particularly tight coupling between ideophones and iconic gestures in everyday speech. The perspective on ideophones developed here has shed light on this relation and suggested a coherent cause for the common coupling. Both ideophone and iconic gesture are holistic representations of complex states of affairs; and both form integral parts of a single multimodal performative depictive act.

The coupling of ideophone and iconic gesture in composite utterances offers a unique opportunity to study iconicity in relation to the affordances of medium and modality. Despite interesting differences in devices of iconicity and degree of conventionalisation, the fundamental kinship of ideophones and iconic gestures as depictions of sensory imagery is clear. With a twist to Kendon’s (1980) classic “Gesticulation and speech: two aspects of the process of utterance”, we may say that ideophone and iconic gesture are two aspects of the process of depiction.

Acknowledgements

I thank the Mawu people and especially Odime Kanairo, Rev. A.Y. Wurapa, and Ruben and Ella Owiafe for their help and hospitality in the field. Portions of this paper were adapted from an unpublished PhD thesis (Dingemanse, 2011a). Nick Enfield and Felix Ameka provided feedback on earlier incarnations of the manuscript, and Adam Kendon provided excellent and incisive comments on a later version. I dedicate this paper to Ruben Owiafe, ideophone performer *par excellence*.

Notes

1. The term ‘mimetics’ is confined mostly to Japanese linguistics, and is usually seen as co-terminous with the more widely used ‘ideophones’ (Akita, 2009; Hamano, 1986). A similarly regionally confined term is ‘expressives’, used in South-East Asian linguistics (Diffloth, 1972; Kruspe, 2004). The most widely used crosslinguistic term for the phenomenon appears to be ‘ideophones’ (Nuckolls, 1999; Voeltz & Kilian-Hatz, 2001).
2. The cartoon was *Canary Row* (Warner Brothers), part of the Sylvester and Tweety Bird series, widely used by David McNeill and many of his students (Kita, 1993, pp. 33–36).
3. This refers to the production of palm oil, in which palm fruit pulp (*sibara*) is wrung in a special type of net (*kàsukutu*) to press out the oil.
4. Zondo’s phrasing illustrates a common inconsistency in discussions of the meaning of different types of signs in composite utterances. The inconsistency has been described by Enfield as follows: “Linguistic items like words are often described merely in terms of what they conventionally *encode* (as standing for lexical *types*), while gestures are typically described in terms of what they non-conventionally *convey* (as standing for utterance-level *tokens* of informative intention)” (Enfield, 2009, pp. 13–14, emphasis in original).

References

- Akita, Kimi (2009). A grammar of sound-symbolic words in Japanese: Theoretical approaches to iconic and lexical properties of Japanese mimetics. PhD dissertation, Kobe University.
- Butterworth, Brian & Uri Hadar (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, 96, 168–174.
- Cassell, Justine & David McNeill (1991). Gesture and the poetics of prose. *Poetics Today*, 12, 375–404.
- Clark, Herbert H. (1996). *Using language*. Cambridge: Cambridge University Press.
- de Ruiter, Jan Peter (2000). The production of gesture and speech. In David McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge: Cambridge University Press.
- Diffloth, Gérard (1972). Notes on expressive meaning. *Chicago Linguistic Society*, 8, 440–447.
- Dingemans, Mark (2011a). The meaning and use of ideophones in Siwu. PhD dissertation, Radboud University, Nijmegen.
- Dingemans, Mark (2011b). Ideophones and the aesthetics of everyday language in a West-African society. *The Senses and Society*, 6, 77–85.
- Dingemans, Mark (2011c). Ezra Pound among the Mawu: Ideophones and iconicity in Siwu. In Pascal Michelucci, Olga Fischer, & Christina Ljungberg (Eds.), *Semblance and signification, iconicity in language and literature* (pp. 39–54). Amsterdam: John Benjamins.
- Dingemans, Mark (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6, 654–672.
- Dingemans, Mark (in press). Expressiveness and system integration. On the typology of ideophones, with special reference to Siwu. *STUF — Language Typology and Universals*.
- Doke, Clement Martyn (1935). *Bantu linguistic terminology*. London: Longmans, Green and Co.

- Ekman, Paul & Wallace Friesen (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. *Semiotica*, 1, 49–98.
- Enfield, Nick J. (2013). Doing fieldwork on the body, language, and communication. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill, & Sedinha Teßendorf (Eds.), *Handbook body — language — communication*. Berlin: Mouton de Gruyter.
- Enfield, Nick J. (2005). The body as a cognitive artifact in kinship representations: Hand gesture diagrams by speakers of Lao. *Current Anthropology*, 46, 51–81.
- Enfield, Nick J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.
- Fortune, George (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.
- Goffman, Erving (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodman, Nelson (1968). *Languages of art: An approach to the theory of symbols*. Indianapolis: Bobbs-Merrill.
- Goodwin, Charles & John Heritage (1990). Conversation analysis. *Annual Review of Anthropology*, 19, 283–307.
- Güldemann, Tom (2008). *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.
- Hamano, Shoko Saito (1986). The sound-symbolic system of Japanese (ideophones, onomatopoeia, expressives, iconicity). PhD dissertation, University of Florida.
- Hetherwick, Alexander (1889). *Introductory handbook of the Yao language*. London: Society for Promoting Christian Knowledge.
- Holler, Judith & Geoffrey Beattie (2003). How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process? *Semiotica*, 146, 81–116.
- Kendon, Adam (1980). Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague: Mouton.
- Kendon, Adam (1997). Gesture. *Annual Review of Anthropology*, 26, 109–128.
- Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam (2008). Signs for language origins. *The Public Journal of Semiotics*, 2, 2–29.
- Kita, Sotaro (1993). Language and thought interface: A study of spontaneous gestures and Japanese mimetics. PhD dissertation, University of Chicago.
- Kita, Sotaro (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics*, 35, 379–415.
- Kita, Sotaro (2001). Semantic schism and interpretive integration in Japanese sentences with a mimetic: Reply to Tsujimura. *Linguistics*, 39, 419–436.
- Klassen, Doreen Helen (1999). You can't have silence with your palms up: Ideophones, gesture, and iconicity in Zimbabwean Shona women's ngano (storysong) performance. PhD dissertation, Indiana University.
- 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 F. K. Erhard Voeltz & Christa Kilian-Hatz (Eds.), *Ideophones* (pp. 183–191). Amsterdam: John Benjamins.

- Levinson, Stephen C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Mandel, Mark (1977). Iconic devices in American Sign Language. In Lynn A. Friedman (Ed.), *On the other hand: New perspectives on American Sign Language* (pp. 57–108). New York: Academic Press.
- McNeill, David (1992). *Hand and mind*. Chicago: University of Chicago Press.
- Noss, Philip A. (2001). Ideas, phones and Gbaya verbal art. In F. K. Erhard Voeltz & Christa Kilian-Hatz (Eds.), *Ideophones* (pp. 259–270). Amsterdam: John Benjamins.
- Nuckolls, Janis B. (1996). *Sounds like life: Sound-symbolic grammar, performance, and cognition in Pastaza Quechua*. New York: Oxford University Press.
- Nuckolls, Janis B. (1999). The case for sound symbolism. *Annual Review of Anthropology*, 28, 225–252.
- 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* (pp. 233–251). Washington, DC: Smithsonian Press.
- Perniss, Pamela, Robin L. Thompson, & Gabriella Vigliocco (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, 1, 1–15.
- Reiter, Sabine (2012). Ideophones in Awetí. PhD thesis, Universität Köln.
- Samarin, William J. (1971). Survey of Bantu ideophones. *African Language Studies*, 12, 130–168.
- Schegloff, Emanuel A. (1984). On some gestures' relation to talk. In John Maxwell Atkinson & John Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266–296). Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In Nick J. Enfield & Stephen C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 70–96). Oxford: Berg.
- Schlegel, J. Bernhard (1857). *Schlüssel zur Ewe-Sprache, dargeboten in den grammatischen Grundzügen des Anlo-Dialekts derselben*. Stuttgart: Valett.
- Stivers, Tanya (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction*, 41, 31–57.
- Streeck, Jürgen (2008). Depicting by gesture. *Gesture*, 8, 285–301.
- Tsujimura, Natsuko (2001). Revisiting the two-dimensional approach to mimetics: A reply to Kita (1997). *Linguistics*, 39, 409–418.
- Voeltz, F. K. Erhard & Christa Kilian-Hatz (Eds.) (2001). *Ideophones: Typological studies in language*. Amsterdam: John Benjamins.
- Werner, Alice (1919). *Introductory sketch of the Bantu languages*. London: Kegan Paul.
- Westermann, Diedrich Hermann (1927). Laut, Ton und Sinn in westafrikanischen Sudansprachen. In Franz Boas (Ed.), *Festschrift Meinhof* (pp. 315–328). Hamburg: Augustin.
- Zondo, Jerry (1982). Some aspects of the ideophone in Ndebele. *Zambezia*, 10, 111–126.

Author's address

Mark Dingemanse
Language & Cognition Group
Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen
The Netherlands
mark.dingemanse@mpi.nl

About the author

Mark Dingemanse is a research staff member of the Max Planck Institute for Psycholinguistics. He has carried out extensive fieldwork in Ghana working on ideophones, iconicity, and social interaction.