Author(s): Tuuri, Kai; Pirhonen, Antti

Title: Gestural expressions in use for unveiling dynamic experience attributed to verbs

Year: 2013

Version:

Please cite the original version:

Tuuri, K., & Pirhonen, A. (2013). Gestural expressions in use for unveiling dynamic experience attributed to verbs. In Proceedings of the 10th International Gesture Workshop and the 3rd Gesture and Speech in Interaction Conference. Tilburg University. http://tiger.uvt.nl/pdf/papers/tuuri.pdf

Gestural expressions in use for unveiling dynamic experience attributed to verbs

Kai Tuuri (kai.tuuri@jyu.fi)
Department of Music, University of Jyväskylä
FI-40014, Finland

Antti Pirhonen (antti.pirhonen@jyu.fi)
Department of Computer Science and Information Systems, University of Jyväskylä
FI-40014, Finland

Abstract

The focus of this paper is on justifying the presented experimental design that aims at examining the enactive linkages between a verb's content and a sensorimotor experience of movement. The experiment utilised spontaneous production of hand and vocal gestures for expressing the energetic feel attributed to a word. Preliminary qualitative analysis of the expressions shows degrees of similarity in terms of experiential movement qualities. These results imply that conceiving a verb's meaning is not necessarily far removed from bodily action.

Keywords: Gestures; vocal gestures; enaction; dynamic experience; verb content; vitality; cross-modality.

Introduction

The role of words is pivotal in studying human experience. Despite the ever-continuing advances in ways of measuring different psycho-physiological responses and neurological processes, the phenomenological essence of our experiences seems to remain beyond the reach of the traditional third-person position of natural sciences. The need for first-person methods for describing one's subjective feelings and perceptions is therefore not going away (e.g., Varela & Shear, 1999). Such descriptions rely heavily on verbalisations, and this aspect raises the question of how the immanent and transient nature of experience itself and the words used for describing it are interconnected.

With respect to verbal reflecting of one's experience of, for example, listening to a musical piece, we have to take into account whether the articulated words are capable of grasping the essential qualities of the dynamic phenomenon of being involved with music. Or perhaps these words refer to some inferred abstractions with only an indirect relationship to the music? In other words, when somebody says that a certain passage of music feels like 'escaping', is there something in the experiential dynamic forms of music that concretely manifests the feeling (e.g., of running scared), or does the idea of escape simply refer to an imaginative creation of one's mind? As will be discussed below, these two explanations need not rule each other out.

Imagined movement in words

George Lakoff and Mark Johnson (1999) have long argued that concepts we use in thinking and talking (i.e., the content of words) ultimately have their foundation in the dynamic, sensorimotor ontology of a living, experiencing body. According to their theory, the root meanings of a concrete sensorimotor level are imaginatively projected in thinking, and abstracted through utilisation of conceptual metaphors. Other authors (Gallese & Lakoff, 2005; Stern, 2010) have also suggested that imagined movement or imagined action might be a central aspect in understanding the meaning of words. This hypothesised, usually preconscious mental effort consists in motor enaction of the type and profile of action/movement that is being attributed to a word. Enaction has underpinnings in a neural simulation of an action that shares the same neural substrate involved in performing those actions. With respect to understanding words, such theories emphasise the enactive experience of action while also stressing its integration with the imaginative processes of human thinking and reasoning. Enactive linkage between language and sensorimotor experience is here assumed to work in a two-way manner: both in concretising and situating the experiential content of an abstract word as well as in translating living, dynamic experiences into the form of words. The attribution to the dynamicity (of concrete movement or action) is obviously most evidently present in verbs.

Aims and implications of the study

This paper is a work-in-progress report of an empirical study in which dynamic experience attributed to Finnish-language verbs is in the spotlight. By utilising spontaneous gestures for expressing the energetic feel experienced in the verbal content, the intention of the present study is in manifesting the embodied experience in observable physical movement in the form of nonverbal hand gestures and vocalisations. It is assumed that the outlined experimental method, utilising spontaneous corporeal articulation (Leman, 2008), taps into the process of imagined activity involved in the process of understanding the word's content. The focus of this examination is not so much on the types of action these gestures mimic as it is on more general dynamic qualities and *forms of vitality* (Stern, 2010) they manifest in their movement. As the vitality forms are supposedly *amodal* in nature, different modalities (i.e., hand and vocal gestures) are expected to manifest qualities of dynamic experience more or less similarly. In all, a certain amount of intra-word consistency of gestural expressions is likely. But instead of underlining differences between the words, the emphasis will be on analysing the similarities and variations in gestural expressions of the same word content, potentially revealing varying qualities in the underlying imagined action. No strong presumptions or expectations about the results are formulated. The agenda of this study is

Table 1: Words chosen for the study translated in English. Original Finnish words are in parenthesis.

	Force-oriented	*Space-oriented*	*Intention/Direction-oriented*
Impulsive effort	bursting (purskahtaen)	opening (aueten)	reaching for (kurottautuen)
	exerting (ponnistaen)	flipping over (kiepahtaen)	dropping oneself (pudottautuen)
	releasing (vapauttaen)	crumpling (rutistuen)	streaking (syöksähtäen)
Continuous effort	fading (hiipuen)	floating (leijuen)	rising (kohoten)
	dragging (raahaten)	suspending (pidättäen)	sidling (hivuttautuen)
	tensing (jännittäen)	stretching (venyttäen)	withdrawing (vetäytyen)
Iterative effort	pulsating (sykkien)	swaying (keinuen)	walking (askeltaen)
	trembling (vavahdellen)	tremoring (täristäen)	groping (tunnustellen)
	stroking (silitellen)	rolling (rullaten)	ramming (juntaten)

motivated by exploration and the analysis will mostly utilise open-ended, qualitative and content-based approaches.

Deeper understanding of the present research subject has many potential implications. First, in general it sheds light on the embodied background of meaning-making involved in verbal descriptions of experiences that relate to, e.g., music listening. Second, it may help us in conceiving the moving forms of, e.g., dance, music and cinema in terms of gestural forms along with the words they potentially correspond with. And third, it can give us clues for choosing words that correspond to certain dynamic experiences. This is useful information, for example, in selecting suitable words to be used as design concepts (or narrative design metatopics, see Hug & Misdariis, 2011) for referring to the intended qualities of use experience in the design of human-technology interaction.

Experiment

Choosing the words

Vitality affects (Stern, 1985) and forms of vitality (Stern, 2010) refer to *contours of feeling* that consist of dynamic changes and fluctuations in our temporal, embodied feeling of being alive – and in a feeling of someone else being alive as well. Rather than concerning categorical feelings or affects, vitality forms aim at describing temporal characteristics related to the changes in a flow of feeling. The words that seem to best refer to such qualities are kinetic expressions such as 'opening', 'accelerating' or 'bursting'.

For this type of experiment, it is important to use the participants' native language and therefore Finnish verbs were chosen for the study. More specifically, verbs in an instructive case were used, which in the Finnish language has the meaning 'by means of', e.g., 'aueten' precisely translates into English as 'by opening'. This verbal form was chosen with the aim of emphasising the dynamicity in a verb's content (i.e., *how* something happens) implicitly as an instruction for the participants to act out their energetic feel attributed to the word. It seems obvious that the results of this study will be language-dependent at least to some degree. Even 'direct' translations of words into a different language might utilise a somewhat different imagined activity as a basis of its meaning.

In order to ensure adequate diversity in the verbs chosen and to avoid choosing the words in a completely arbitrary manner, a two-dimensional framework was first formulated for profiling potentially different characteristics in dynamic experience attributed to verbs. Both dimensions concern the kinaesthetic *effort* implied in the word's meaning. The first dimension adopts Pierre Schaeffer's (Schaeffer & Reibel, 1998) typological system of defining energetic envelopes of gesture production as either (1) impulsive/punctual, (2) sustained/continuous or (3) iterative (see also Godøy, 2006). The second dimension adopts Stern's (2010) notion of five dynamic aspects in an experience of vitality, and (out of these) picks up the following as relevant for characterising the orientation of an energetic effort: (1) force, (2) space and (3) intention/directionality. The two dimensions were combined together to form a 3×3 matrix yielding a profile of nine different types of dynamicity. Three verbs were chosen to appropriately represent each of the nine types, resulting in 27 words altogether. The chosen Finnish words and their English translations are shown in Table 1. The total number of words is a compromise: a relatively large amount of words should give a better general overview of the studied phenomenon. But of course, a smaller amount of words would have given more opportunities for studying the experience behind each word, e.g., through extended interviews.

Participants, setting and procedure

A group of 14 Finnish-speaking students of the University of Jyväskylä were recruited to participate in the experiment. The average age in the group was 24.1 years (with an SD of 3.6 years). The number of male and female participants was equal (7). Most of the participants (78%) were students of musicology, while two of the participants studied music education and one studied a subject other than music (romance philology). Two of the participants were postgraduate students. All participants were rewarded with a movie ticket.

The experiment consisted of two gesture-production tasks (one using hand gestures and the other using vocal gestures) and a short interview about each participant's experience of the tasks. In the production tasks participants were asked to spontaneously gesticulate/express the energetic feel that

was instantly evoked in them after reading a word. The instructions emphasised that the execution of gestures should be based on intuition and an instant feeling rather than on rational reasoning of any kind. The experimental setting, with a participant and a researcher positioned almost face-to-face (see Figure 2), was intended to strengthen the person-to-person communicative validity of gestural expressions. Such bodily expressions may be more validly taken as descriptions of participants' subjective experience in a *second-person* position, i.e., experience as corporeally articulated 'from me to you' (Leman, 2008; Varela & Shear, 1999). However, in this experiment the face-to-face setting also required that the researcher constantly and self-consciously avoided offering any additional stimuli to a participant.

In the hands-condition task participants expressed each verb's dynamic content by using their hands (with the rest of the upper body when felt appropriate). The voice-condition task was otherwise similar but instead of hands it utilised participants' vocal apparatus for nonverbally expressing the dynamic content. Half of the participants did the tasks in reverse order, in order to counterbalance the learning effect between the tasks. The cross-modal design of the experiment allows the examination of possible interconnections between the two embodied ways of expressing principally the same mental idea (the dynamicity or imagined movement in a verb). From a musicological perspective, voice-condition is also interesting because the resulting gestures are produced in an acoustic form and thus may be compared more directly to the forms of music.

In both tasks, sequences of words were presented one word at a time in random order on a display in front of the seated participant. The displaying of each word was preceded (3 seconds) by a short sound signal so that the participant could orientate herself to reading the upcoming word, which stayed on the screen a few seconds. The participant started executing the gestural expression immediately after reading the word. The researcher, who was present in the experiment, controlled the presentation of words and allocated enough time for performing each gesture (expressions typically lasted some few seconds). The overall procedure of a task started with three rehearsal words for accustoming oneself to the production task. After the warm-up, a randomised sequence of 27 words was completed twice in both tasks. In all, two gestural expressions were produced for each word in both conditions (with hands and with voice).

The experimental tasks were recorded with two video cameras and a sound recorder. The main camera in front of a participant (see Figure 1) used a high definition 1080p50 format (with a double frame rate and increased resolution for capturing motion) while the secondary camera covered an overall view of the setting from the side (Figure 2). A highly directional AKG CK98 microphone was used for sound recording. Participants wore gloves of a different colour on each hand, thus providing an option for tracing the hand trajectories separately in video motion analysis (see, e.g., Jensenius, 2012).

Figure 1: A participant expressing the verb 'bursting' viewed from the main camera.

Figure 2: An overview of the experimental setting in a situation shown in Figure 1.

Preliminary results

As the examination of the material gathered from the experiment is still at early stage, this section aims at providing a preliminary 'snippet' of the ongoing analysis and results. In general, the experiment was successful in terms of gaining video and audio recordings of participants' gestural expressions in an expected manner. According to the interviews, participants mostly felt that both tasks or at least one or other of the tasks were easy to perform. Even the participants who found tasks difficult managed to produce expressions in a relatively spontaneous and rapid manner. From all of the experimental trials, only a few times did a participant fail to produce an expression for a word. Because all of the rare failures occurred within a task's first sequence of words, and because participants were more satisfied with their performance in a second trial of the word, the following examination is limited to the gestures performed in the latter sequence of each task. Moreover, for this preliminary analysis, only two words, 'bursting' (purskahtaen) and 'floating' (leijuen), out of the 27 were chosen as targets of examination.

In order to better understand the material and see what kind of further analysis would be appropriate, we started by profiling the gestural samples in terms of their similarity in their dynamic qualities within a word. This is achieved by performing a qualitative evaluation in the manner of paired comparisons between the samples of different participants. In such analysis, all 14 samples produced for a word in a condi-

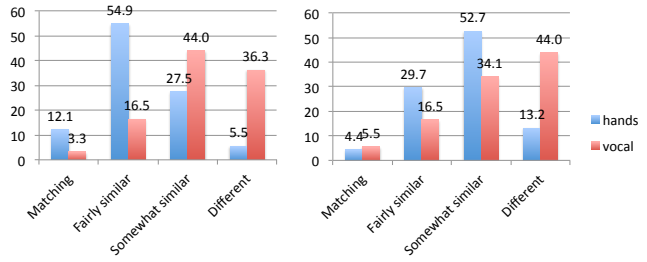

Figure 3: The percentage distributions of similarity assessments of paired comparison tests (n=91 each). Results for the word 'bursting' are on the left and for 'floating' on the right.

tion are subjectively compared to each other pair-by-pair and classified by their similarity in terms of movement qualities (i.e., not primarily in terms of the objective shape of movement). Each of the 91 pairs were assessed by the researchers using a 4-point Likert scale corresponding to 1) matching 2) fairly similar 3) somewhat similar or 4) different dynamicity, resulting in a matrix that denotes similarity relations between samples. The distributions of the assessment results are shown in Figure 3. Goodness-of-fit tests were performed to determine the dependency of word or condition on the similarity assessments in general. As the histograms in Figure 3 also imply, the ways the experienced similarity is distributed across gestural samples differ significantly by the word, $\chi^2(3, N{=}364) = 57.65$, $p{<}.001$, and by the condition, $\chi^2(3, N{=}364) = 9.57$, $p{<}.05$.

Cross-modal comparisons (across conditions) will be included in the succeeding analysis of similarity in order to shed more light on the inter-condition relations of samples. In all, the most important function of the comparisons is to outline such gestural expressions (e.g., groups of samples with extreme or inconsistent similarity relations to each other) that could potentially reveal more detailed characteristics of the studied phenomenon in the analysis to follow.

Conclusions and future directions

The focus of this work-in-progress report was in justifying the presented experimental design for unveiling the dynamic experience attributed to words. The tentative results of the experiment were encouraging: the easiness and intuitiveness of gesturally expressing the energetic feel attributed to words points to the prospective involvement of an enactive linkage between sensorimotor experience and a word's content. Despite the expected variation across hand and vocal gestures for the same content, the expressions had similarities in their felt dynamicity. However, further analyses are needed for closer examination of these similarities. These will include content analysis of interviews and the recorded commentaries relating to paired comparison analysis. Phenomenological analysis utilising, e.g., an empathetic approach (Tarvainen, 2012) on the enaction process may also be utilised for a smaller number of chosen samples. Other possible analysis may include quantitative extraction of motion data and qualitative analysis of visualised motion contours.

We hope that this type of research could partly contribute to developing a framework for studying human experience that would have potential for building a bridge between first-person and third-person approaches. Even the preliminary results presented in this paper suggest that verbal expressions are not necessarily that far removed from a bodily ontology of concrete activity. Other authors have made similar observations. Within the context of music listening, for example, Rautio (2007) and Tarvainen (2012) have outlined linkages between embodied qualities of movement in music and their metaphors in verbalisations. From such focal point, it is also possible to extend the examination to the physical forms of music and musical performance.

Acknowledgments

This work is funded by the Finnish Funding Agency for Technology and Innovation (project diary num. 313/31/12).

References

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, *22*(3), 455–479.

Godøy, R. (2006). Gestural-sonorous objects: embodied extensions of schaeffer's conceptual apparatus. *Organised Sound*, *11*(2), 149.

Hug, D., & Misdariis, N. (2011). Towards a conceptual framework to integrate designerly and scientific sound design methods. In *Proceedings of the 6th Audio Mostly: A conference on interaction with sound* (pp. 23–30).

Jensenius, A. R. (2012). Evaluating how different video features influence the visual quality of resultant motiongrams. In *Proceedings of the 9th sound and music computing conference.* Copenhagen, Denmark.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.

Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: The MIT Press.

Rautio, R. (2007). Metaphors of motion in listeners' verbal reports. In *Proceedings of the 10th international conference on music perception and cognition.* Sapporo, Japan.

Schaeffer, P., & Reibel, G. (1998). *Solfege de l'objet sonore*. Paris: INA/GRM (original work published in 1967).

Stern, D. (1985). *The interpersonal world of the infant*. New York, NY: Basic Books.

Stern, D. (2010). *Forms of vitality: exploring dynamic experience in psychology, the arts, psychotherapy, and development*. Oxford Univ Pr.

Tarvainen, A. (2012). *Laulajan ääni ja ilmaisu-kehollinen lähestymistapa laulajan kuuntelemiseen*. Tampere University Press.

Varela, F., & Shear, J. (1999). First-person methodologies: What, why, how. *Journal of Consciousness studies*, *6*(2-3), 1–14.