ELŻBIETA GÓRSKA
University of Warsaw
elzbieta.gorska@uw.edu.pl

TEXT-IMAGE RELATIONS IN CARTOONS. A CASE STUDY OF IMAGE SCHEMATIC METAPHORS

Abstract

As compared to their purely verbal manifestations, multimodal realizations of image schematic metaphors have received far too little attention in cognitive linguistics than they would deserve. It will be argued that image schemas (Johnson 1987, Talmy 1988), since they are skeletal conceptual structure, afford an excellent source domain for metaphors that are realized verbo-visually in cartoons. The cartoons selected for this study are all by Janusz Kapusta, a Polish artist, whose works have appeared every week in the Polish magazine "Plus-Minus" for over ten years. In contrast to the gestural medium, films and music, where the relevant elements of image schematic source domains of metaphor are never fully available at once, the cartoons give a "snapshot" of a conceptual image which is ready for inspection as a single Gestalt. They are therefore a good testing ground for discussing the question of how the visual and the verbal modality interact in spatialization of abstract ideas. Providing insights into the function of multimodal metaphors and levels of their activation, the discussion contributes to the ongoing debate on the conceptual nature of metaphor and the embodiment of meaning. The results of the study are also considered in relation to the role of verbo-pictorial metaphor in structuring abstract concepts in a creative way.

1. Introduction

In studies of text-image relations, the issue of establishing which elements of the verbal message correspond to the elements of the visual message is one of the main queries. Equally important is the question of how to evaluate the relative contribution of each of the two modes to the multimodal communicative act as a whole.

As Bateman (2014) shows, the quickly developing field of text-image studies, while offering a range of approaches is still at a stage of searching for some well-founded tools of analysis. This study will, hopefully, contribute to this search. It will also shed some light on multimodal metaphor research which has already gained firm ground in cognitive linguistics.

I will focus here on verbal and pictorial manifestations of image schemas which, in terms of their original conception (Johnson 1987; Lakoff 1987), are prelinguistic patterns of sensory-motor experience that emerge from (or are grounded in) our bodily interactions with the environment, of which bodily movement through space, perceptual interactions, and manipulation of objects play a crucial role (for subsequent additions to the original list of image schemas see Hampe 2005b and Mandler 2010; cf. also Johnson 2007). They characterize both states and processes and, what is crucial for our concern, are "cross-modal", i.e. are able to transfer information between different sensory systems. Immediately relevant to this study is Johnson's groundbreaking claim that image schemas are commonly recruited as source domains of conceptual metaphors. However, as compared to purely verbal manifestations of image schemas, their multimodal realizations have received far too little attention in cognitive linguistics than they would deserve. Still the diversity of the relevant research is remarkable. Suffice it to mention here studies on co-speech gesture by Cienki (2005, 2013), Calbris (2008), Mittelberg (2010); on film by Forceville (2006, 2013); on text painting in music by Zbikowski (2000); and on commercial brands and logos by Pérez Hernández (2013, 2014) (see also Forceville, Urios-Aparisi 2009b; Pinar Sanz 2013; Górska 2014a). This study will aim to show that image schemas function as a source domain of multimodal metaphors whose construction seems pivotal to a coherent interpretation of a number of cartoons by a Polish artist Janusz Kapusta.[1] Following Forceville and Urios-Aparisi (2009a: 4) I assume that a defining characteristic of a multimodal metaphor is that its "target and source are rendered exclusively or predominantly in two different modes/modalities".[2]

The cartoons selected for this study were originally published in a Polish weekly "Plus-Minus", yet my data sample comes from a recently published book by Janusz Kapusta (2014) which marked the 10th anniversary of his weekly collaborations with this magazine. It comprises over 200 cartoons, out of about 500 published during this ten year period. Let me add that Kapusta has been living in New York since 1981, and since then his work has been also regularly published in, among others, "The New York Times", "The Wall Street Journal" and "The Washington Post".

It needs to be noted also that since only the image schematic structure of the cartoons is in my focus, when considering the drawings I will not go into matters of design or composition which, of course, would be relevant in a more comprehensive analysis. Such an analysis should also give an account of the dynamics of meaning

---

[1]   Bateman (2014: 180) devotes only a short section to image schemas and metaphor and, following Kennedy (1982) and Forceville (2005), observes that they "also may motivate certain conventions of pictorial depiction".

[2]   The terms *mode* and *modality*, as is common in the literature on the subject, are used here interchangeably (see Forceville and Urios-Aparisi 2009a: 4).

construction and the creation of emergent concepts in particular. For the present purpose, suffice it to add that, in terms of the multispace model of the Blending Theory (Fauconnier, Turner 1998, 2002), the discussed image schemas would capture the structure and organization of the verbal and/or the visual inputs, and define numerous correspondences between them.

2. Image-schematic metaphors in Janusz Kapusta's cartoons

Let us first consider some relevant aspects of the cartoons' structure. The cover of the book from which my data comes may serve as an illustration.

The cartoons always have a protoganist whose shape, as the cover clearly indicates, reminds of a Buddha or a chess pawn who formulates a caption that is, as a rule, represented above the protagonist's head and linked to it by a short line segement (on the cover it points to the title "Plus Minus"). As an enlightened sage (a Buddha) or an everyman (a pawn), the protagonist dwells upon a whole range of issues, including various aspects of human life and human condition, religion, politics and the world and universe at large. Functionally, the captions may be regarded as "speech balloons" or "thought bubbles", with the little line segment playing a role of a semiotic tool for interfacing the verbal and the pictorial.[3] As in other multimodal genres of this kind, the very fact that the verbally and the pictorially expressed messages co-occur in a single communicative act requires that they be integrated together into a unified conceptual whole. In building this conception the recipient is thus confronted with a problem solving task – a multimodal riddle whose solution itself might be a pleasurable reward.

---

[3]   In the literature, the structure of the interface is described in terms of the carrier-tail-root configuration, and the verbally expressed information (typically the carrier) and pictorially expressed information (the root) as well as the tail, which connects root and the carrier, may appear "in a variety of visual forms" (Bateman 2014: 109; see also Forceville et al. 2010).

Considering first the caption in Example 1, which reads „Można zrobić wiele błędów, ale nie warto się pomylić" – 'You can make many errors, but it's not worth to be mistaken' (Kapusta 2014: 208),[4] we may observe that, except for the metaphorical idiom *robić błędy* 'make errors', which is based on the ontological metaphor IDEAS ARE OBJECTS, the interpretation of the caption itself does not rest upon the activation of any other image schematic metaphor.

Example 1: 'You can make many errors, but it's not worth to be mistaken' (Kapusta 2014: 208, transl. E.G.)

Rather, it is the need to integrate the verbal and the pictorial modality into a single coherent message that prompts the construction of multimodal image schematic metaphors, with the visual modality providing the relevant source domain and the verbal modality – the relevant target concepts. Apart from the OBJECT schema, four other image schemas are relevant here: STRAIGHT, PATH, FORCE and LONG-SHORT (the latter is one of ATTRIBUTE schemas).[5] And in particular, this image schema complex provides motivation for establishing a cohesive tie between the long straight arrow going through the protagonist's head with the idea of 'not being mistaken (possibly in life)'.[6] For the sake of the present discussion, this creative metaphor can be phrased as: NOT BEING MISTAKEN IN LIFE IS SELF-PROPELLED MOTION FORWARD ALONG A SINGLE STRAIGHT PATH. Crucially, the construction of this creative metaphor is well-founded in our conventional metaphorical thought. On the one hand, it is motivated by our cultural model in terms of which, as Cienki puts it, "normal

[4]   I'm very grateful to Janusz Kapusta for granting to me the permission to reprint the cartoons that I consider in this study.

[5]   For image schema groupings see, in particular, Cienki (1997). Note also that ATTRIBUTE schemas are not included in Johnson's original (1987) list of image schemas; however, they have been commonly discussed and also applied in research on image schematic metaphors in user interface design (see, in particular, Hurtienne 2014: 319, and the literature cited therein; cf. also Macaranas et al. 2012).

[6]   For multimodal cohesive ties see Bateman (2014: 167).

(default, 'good') action, speech, and thought are characterized metaphorically as motion along a straight path" and, on the other hand, by a metaphor coherent with this model, namely CORRECT IS STRAIGHT (Cienki 1998: 125). And, by implication, the notion of being wrong or incorrect is conceived of in terms of straying or wandering off a straight path: WRONG IS NOT STRAIGHT (Cienki 1998: 125). This conventional metaphor quite clearly underlies the construction of another multimodal metaphor which helps in establishing cohesive ties between the text and the drawing: the verbally expressed target concept of 'making errors' can be integrated with the many short line vectors in the drawing. Moreover, the fact that the majority of these vectors are shaped as bent, curved, twisted or turned is motivated by the interaction of the STRAIGHT image schema with the image schemas for forced movement (Talmy 1988, cf. in particular, a change in the directionality of the force vector is understood in terms of caused motion). And how can we account for the four short straight line vectors in the centre of the drawing? It seems that the relevant aspects are their shortness and divergent directionality – they can be conceived of as transient departures from a continuous stable motion along a straight and long path of 'not being mistaken'. We may note also that in this example the level of image schemas activation is very low – except for the OBJECT schema, which inheres in the conceptions of errors as objects (cf. *robić błędy* 'make errors'), they are only cued visually.

To analyze the next cartoon, let us first consider the drawing.
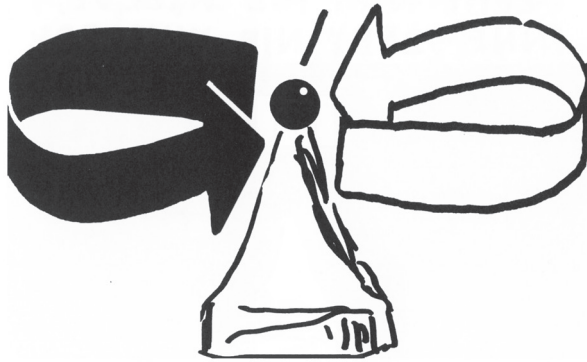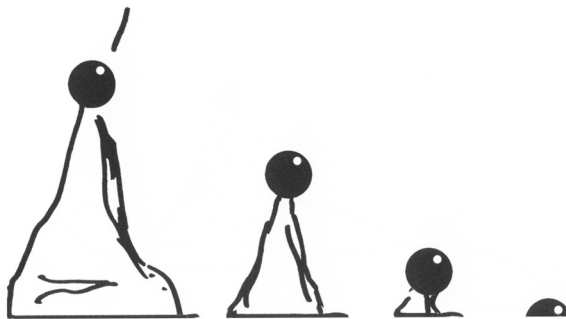

Example 2: Kapusta (2014: 218)

The pictorial mode itself suggests that the protagonist is affected by something coming from behind and something approaching from the front. This interpretation would require an activation of an image schema complex comprising the OBJECT, PATH, MOTION, FRONT-BACK, and FORCE image schemas. One may also assign an image schematic metaphorical interpretation to the conspicuous black vs. white attribute of each of the two vectors, with black (and dark) metaphorically associated in our culture with something negative and white (and light) with something positive.

When we link this image schematic structure with the caption saying „Niespełnione marzenia dręczą nie mniej niż koszmary przeszłości" – 'Unfulfilled dreams torment you no less than the nightmares of the past' (Kapusta 2014: 218,

transl. E.G.) we can now see that the verbal mode introduces the target concepts which are cued in the visual modality by the two vectors – the nightmares, i.e. the past events are – in all likelihood – approaching the ego from behind and the yet unfulfilled dreams are moving towards the ego from the front. What underlies this conception is the deictic MOVING TIME model (Lakoff 1993; Lakoff, Johnson 1999: 139–150; Evans, Green 2006: 84–87; see also Lewandowska-Tomaszczyk 2016) in an extended variant: the future events – our yet unfulfilled dreams are moving towards the ego from the front, which is in accord with the conventional Moving Time Model, but unlike in this model, also the past events – the nightmares of the past – are moving towards the ego from behind. For our purpose, the creative metaphors which make this communicative act coherent can be rendered as: (AS YET) UNFULFILLED DREAMS ARE FORCES MOVING TOWARDS THE EGO FROM THE FRONT and NIGHTMARES (OF THE PAST) ARE FORCES MOVING TOWARDS THE EGO FROM BEHIND. A careful observer may notice also that the fact that the two arrows are of the same size matters – it may be interpreted as a visual analog of the verbal description of the same intensity level with which 'nightmares of the past' and 'as yet unfulfilled dreams' affect the protagonist (cf. the expression *nie mniej* 'no less than'). As for the image schema activation, only OBJECT and FORCE are cued in the two modes (cf. the objectified *marzenia* 'dreams' and *koszmary* 'nightmares' and the verb *dręczyć* 'torment' in the verbal mode and the corresponding arrows – the force vectors – encircling the protagonist in the drawing), and all the remaining schemas, namely the PATH, MOTION, and ATTRIBUTE (BLACK/DARK-WHITE/LIGHT) are cued visually only, hence the degree of overlap is still quite low.

Example 3 illustrates a more complex interplay of the two modes. To see this let us begin with the caption saying „Przemijamy przez całe życie, ale pod koniec jakby bardziej" – 'We are passing throughout the whole life but, when close to the end, as if a bit more' (Kapusta 2014: 70). The first clause is based on the conventional metaphor LIFE IS A JOURNEY with the verb *przemijać* 'elapse, pass, go by (in time)'[7] providing

PRZEMIJAMY PRZEZ CAŁE ŻYCIE,
ALE POD KONIEC JAKBY BARDZIEJ

Example 3: 'We are passing throughout the whole life but, when close to the end,
            as if a bit more' (Kapusta 2014: 70, transl. E.G.)

access to the MOVING EGO model of TIME, and thereby to the MOTION schema which lies at its core. When we take a look at example 3 we can establish a cohesive tie between the verbally expressed conception of our passing throughout our life-span and the PATH schema that is realized in the drawing as a line formed by a series of four figural shapes of decreasing size. Since the correlation of the PATH and MOTION schemas recurs regularly in our embodied experience, the interpretation of the cartoon in terms of this cross-modal cohesive tie seems natural and well-motivated.

Integration of this conception with the second clause crucially relies on a cross-modal activation of the SCALE schema: the verbal mode specifies that the intensity of 'our passing throughout the life-span' increases along some scale, and the pictorial mode, relying on the UP/DOWN and BIG/SMALL schemas, depicts this intensity scale in terms of the protagonist's dramatic change in form and size along the vertical dimension. For the sake of this discussion, the creative multimodal metaphor which underlies this interpretation can be phrased as: PASSING IN LIFE IS MOTION FORWARD ALONG A HORIZONTAL PATH COINCIDING WITH THE MOTION DOWNWARD ALONG THE VERTICAL PATH OF AN EGO DISAPPEARING IN THE GROUND. Importantly, the embodied inferential structure of the UP/DOWN and BIG/SMALL schemas triggers the assignment of axiological values to the intensity scale – with its endpoints – the initial state of 'being up and big' (hence active, important and in the peak of power)[8] evaluated as positive, and the final state of 'the reduced body being down – almost completely submerged in the ground' (hence, unable to act, unimportant, and powerless) – as negative.[9]

The last example illustrates a high degree of both visualization and verbalization of image schemas, and in the verbal mode it expresses a simile (by means of the adverb *jak* 'like').

CZŁOWIEK JAK BUTELKA – MA ZNACZENIE,
DOPÓKI MA COŚ W SOBIE

Example 4: 'Man like a bottle – has significance till there is something inside' [lit. …
till he/it has something inside] (Kapusta 2014: 199, transl. E.G.)

---

7   Note that this verb has only a temporal sense in Polish.
8   For metaphorical conceptions of life-cycle see Lakoff, Turner (1989: 87–89).
9   For axiology of UP/DOWN see Hampe (2005a), see also Górska (2014a).

In the construction of this multimodal simile, the source domain of the CON-TAINER schema is expressed verbally as *butelka* 'bottle' and pictorially – by the protagonist's body shaped as a bottle with a particular level of liquid inside. The target concept – the notion of significance of a human being – is cued only in the verbal mode via the expression *mieć znaczenie* 'have significance', however it is the pictorial mode that cues a high degree of significance by depicting a high level of the liquid inside the protagonist's body (MORE IS UP). Note also that the mapping of the inferential structure of the UP/DOWN schema onto this conception triggers a positive evaluation of the high degree of significance (POSITIVE IS UP). The thus constructed creative metaphor can be formulated as: THE SIGNIFICANCE OF A HUMAN BEING IS THE CONTENT OF THE CONTAINER. Referring to the dynamic approach to metaphor (Müller 2008), we can conclude that the level of activation of this metaphor is very high for two reasons. First, there is a high degree of cross-modal overlap of the conceptual content, and, second, the presence of the comparative adverb *jak* 'like' draws attention of the recipient to a metaphorical interpretation of this multimodal communicative act.[10]

3. Conclusions

Corroborating multimodal research on image schemas (see, in particular, Cienki 2005, 2013; Calbris 2008; Mittelberg 2010; Forceville 2006, 2013; Zbikowski 2000; Pérez Hernández 2013, 2014), this study has shown that they afford an excellent source domain for metaphors that are realized verbo-visually in Janusz Kapusta's cartoons. Taking the perspective of the cartoons' addressee we can thus say that image schemas constitute the experientially derived means that we all have at our disposal to solve the conceptual riddles that the cartoons pose. Since the creative image schematic metaphors that lie at the core of those riddles, are deeply rooted in our bodily experience they may be easy to interpret by the audience at large.[11] Their experiential motivation also provides a firm ground for their rhetorical functions of attracting attention, improving understanding, inspiring imagination, and rendering the communicative act memorable (see Stöckl 1997 qtd. in Bateman 2014: 230–231).

Viewed in the context of current debates on metaphor in cognitive linguistics (see e.g. Kövecses 2008; Gibbs 2015ab) this study provides supporting evidence for a dynamic view of metaphor (Müller 2008), since the diverse interplay of the two modes – from complementarity to almost complete overlap in multimodal metaphors considered in this study, points to a dynamic activation of metaphoricity.

---

[10] In Müller's approach the number of "metaphoricity indicators" is symptomatic of the level of metaphor activation (or the degree of metaphoricity; see Müller 2008: 198–201; Górska 2014b).

[11] Research on user interface design has provided ample evidence showing that image schemas and their metaphorical extensions help in designing products that are intuitive in use. Specifically, "[i]t was (…) found that task times and error rates are lower when using a version of a product instantiating image-schematic metaphors than when using traditional versions of the same product" (Hurtienne 2014: 320–321, and the literature cited therein).

Last but not least, with respect to recent developments in text-image studies (Bateman 2014), two advantages of the approach based on the analysis of image schematic structure in a multimodal artefact should be mentioned. On the one hand, it allows for establishing cross-modal cohesive ties which are experientially motivated and embodied. Moreover, since in this approach the choice of cross-modally cued elements which are analyzed as a unit is guided by image schemas and their inferential structure the proposed analyses stand a good chance of including in their scope aspects of multimodal construction which might be easily omitted in other approaches.

References

Bateman J.A. 2014. *Text and image. A critical introduction to the visual/verbal divide*. London, New York.

Calbris G. 2008. From left to right…: Coverbal gestures and their symbolic use of space. – Cienki A., Müller C. (eds.). *Metaphor and gesture*. Amsterdam, Philadelphia: 27–54.

Cienki A. 1997. Some properties and groupings of image schemas. – Verspoor M., Lee K.D., Sweetser E. (eds.). *Lexical and syntactic constructions and the construction of meaning*. Amsterdam, Philadelphia: 3–15.

Cienki A. 1998. Straight: an image schema and its metaphorical extensions. – *Cognitive Linguistics* 9: 107–149.

Cienki A. 2005. Image schemas and gesture. – Hampe (2005b): 412–444.

Cienki A. 2013. Image schemas and mimetic schemas in cognitive linguistics and gesture studies. – *Review of Cognitive Linguistics* 11.2: 417–432.

Evans V., Green M. 2006. *Cognitive linguistics: An introduction*. Edinburgh.

Fauconnier G., Turner M. 1998. Conceptual integration networks. – *Cognitive Science* 22: 133–187.

Fauconnier G., Turner M. 2002. *The way we think. Conceptual blending and the mind's hidden complexities*. New York.

Forceville Ch. 2005. Visual representations of the idealized cognitive model of anger in the Asterix album *La Zizanie*. – *Journal of Pragomatics* 37.1: 69–88.

Forceville Ch. 2006. The source-path-goal schema in the autobiographical journey documentary: McElwee, Van der Keuken, Cole. – *The New Review of Film and Television Studies* 4.3: 241–261.

Forceville Ch. 2013. Metaphor and symbol. Searching for one's identity is looking for a home in animation film. – *Review of Cognitive Linguistics* 11.2: 250–268.

Forceville Ch., Urios-Aparisi E. 2009a. Introduction. – Forceville Ch., Urios-Aparisi E. (2009b): 3–17.

Forceville Ch., Urios-Aparisi E. (eds.). 2009b. *Multimodal metaphor*. Berlin, New York.

Forceville Ch., Veale T., Feyaerts K. 2010. Balloonics: The visuals of balloons in comics. – Goggin J., Hassler-Forest D. (eds.). *The rise and reason of comics and graphic literature: Critical essays on the form*. Jefferson (NC), London (UK): 56–73.

Gibbs R.W. 2015a. Does deliberate metaphor theory have a future? – *Journal of Pragmatics* 90: 73–76.

Gibbs R.W. 2015b. Do pragmatic signals affect conventional metaphor understanding? A failed test of deliberate metaphor theory. – *Journal of Pragmatics* 90: 77–87.

Górska E. 2014a. The up/down orientation in language and music. – Brenzinger M., Kraska-Szlenk I. (eds.). *The body in language. Comparative studies of linguistic embodiment*. [= *Brill's Studies in Language, Cognition and Culture* vol. 8]. Leiden: 177–195.

Górska E. 2014b. Dynamiczne podejście do metafory. – *Prace Filologiczne* 64.2: 109–122.

Hampe B. 2005a. When *down* is not bad and *up* is not good enough: A usage-based assessment of the plus-minus parameter in image-schema theory. – *Cognitive Linguistics* 16.1: 81–112.

Hampe B. (ed.). 2005b. *From perception to meaning. Image schemas in cognitive linguistics*. Berlin.

Hurtienne J. 2014. Non-linguistic applications of cognitive linguistics: On the usefulness of image-schematic metaphors in user interface design. – Littlemore J., Taylor J.R. (eds.). *The Bloomsbury companion to cognitive linguistics*. London: 301–324.

Johnson M. 1987. *The body in the mind: The bodily basis of imagination, reason, and meaning*. Chicago.

Johnson M. 2007. *The meaning of the body. Aesthetics of human understanding*. Chicago.

Kennedy J.M. 1982. Metaphor in pictures. – *Perception* 11.5: 589–605.

Kövecses Z. 2008. Conceptual metaphor theory. Some criticisms and alternative proposals. – *Annual Review of Cognitive Linguistics* 6: 168–184.

Lakoff G. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago.

Lakoff G. 1993. The contemporary theory of metaphor. – Ortony A. (ed.). *Metaphor and thought*. [2nd edition]. Cambridge (UK): 202–251.

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

Lakoff G., Turner M. 1989. *More than cool reason. A field guide to poetic metaphor*. Chicago.

Lewandowska-Tomaszczyk B. (ed.). 2016. *Conceptualizations of time.* Amsterdam, Philadelphia.

Macaranas A., Antle A.N., Riecke B.E. 2012. Bridging the gap: Attribute and spatial metaphors for tangible interface design. – *Proceedings of the 6th International Conference on Tangible and Embedded Interaction*. Kingston (Ontario, Canada): 161–168.

Mandler J.M. 2010. The spatial foundations of the conceptual system. – *Language and Cognition* 2.1: 21–44.

Mittelberg I. 2010. Geometric and image-schematic patterns in gesture space. – Evans V., Chilton P. (eds.). *Language, cognition, and space: The state of the art and new directions*. London: 351–385.

Müller C. 2008. *Metaphors dead and alive, sleeping and waking. A dynamic view*. Chicago.

Pérez Hernández L. 2013. Approaching the utopia of a global brand. The relevance of image schemas as multimodal resources for the branding industry. – *Review of Cognitive Linguistics* 11.2: 285–302.

Pérez Hernández L. 2014. Cognitive grounding for cross-cultural commercial communication. – *Cognitive Linguistics* 25.2: 203–248.

Pinar Sanz M.J. (ed.). 2013. *Multimodality and cognitive linguistics*. [Special issue of *Review of Cognitive Linguistics* 11.2]. Amsterdam.

Stöckl H. 1997. *Textstill und Semiotik englischsprachiger Anzeigenwerbung*. Frankfurt am Mein.

Talmy L. 1988. Force dynamics in language and cognition. – *Cognitive Science* 12: 49–100.

Zbikowski L.M. 2000. *Des Herzraums Abschied*: Mark Johnson's theory of embodied knowledge and music theory. – *Theory and Practice* 22.23: 1–16.

Data source

Kapusta J. 2014. *Plus-Minus. Podręcznik do myślenia*. Poznań.