
Joost Schilperoord  
Corresponding Author  
Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, the Netherlands  
j.schilperoord@tilburguniversity.edu

Neil Cohn | ORCID: 0000-0003-0074-1995  
Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, the Netherlands  
n.Cohn@tilburguniversity.edu

Abstract

This paper adopts a construction-grammar approach to multimodal meaning. We provide a detailed analysis of the Before-After-construction used frequently in advertisements, cartoons and Internet memes. We demonstrate that parts of its generic ‘caused-change’ meaning is compositional, and rendered independently from what is overtly expressed by concrete instances of the pattern. The latter hence build on an abstract multimodal construction whose form elements are paired idiosyncratically with meaning, just like linguistic constructions proper. We show that non-standard instances of the Before-After-construction represent deviations based on a systematized standard Before-After-construction. Finally, we argue that the Before-After-construction belongs to a broader inheritance hierarchy of two-image multimodal construction types, while also providing one amongst several options to convey caused-change. Altogether, we demonstrate that multimodal expressions instantiate similar properties as unimodal expressions both across form and meaning.
Keywords


1 Introduction

Most communication that we experience is multimodal, combining spoken or written language with gestures and/or pictures. One type of recognizable multimodal expression is found across advertisements, cartoons, and memes, where two pictures of people or objects are juxtaposed with labels reading before and after, as in the examples in Figure 1. Here, we see an advertisement promoting an anti-aging face cream (1a), a meme showing Barack Obama's face ‘before kids’ and ‘after kids’ (1b), and an editorial cartoon showing Obama before and after the ‘burden of command’ (1c). In all cases, a common pattern emerges where the two images show different states, with the implication that something (crème, kids, the burden of command) has caused the change between states. We will argue that this common pattern of ‘caused change’ reflects a multimodal construction, one that carries partially compositional

Figure 1 Before-After constructions showing a) an advertisement for anti-aging crème, b) a meme about the effects of having kids, and c) a cartoon about the burden of command for the U.S. president.
meaning and builds on idiosyncratic form-meaning mappings, similar to linguistic constructions proper (cf. Fillmore, 1988; Goldberg & Jackendoff, 2004; Jackendoff, 2002).

Research in multimodality is primarily concerned with how multimodal meaning ‘arises’ from combining and integrating modality-specific semantic contributions (cf. Bateman, 2014; Bateman, Wildfeuer & Hiippala, 2017; Forceville, 2009; Martinec & Salway, 2005; O’Halloran, 2008; Zima & Bergs, 2017). In this paper we explore a further view, i.e. that interactions between modalities within a linguistic framework not only create emergent meanings, but also that such interactions might create entrenched encodings similar to unimodal lexical items. Indeed, such constructional properties have been identified about multimodal memes from social media (Dancygier & Vandelanotte, 2016), and co-speech gestures (Ladewig, 2020; Langacker, 2010; Lanwer, 2017; Steen & Turner, 2013). In the case of Figure 1, we will argue that these examples instantiate a Before-After Construction (henceforth, “BA-construction”) stored in long term memory. To argue its constructional nature, we detail the pattern’s formal and conceptual properties and show that these properties not only allow for a variety of productive usages but also link the pattern to an abstract inheritance hierarchy of multimodal construction types.

If it is the case that multimodal constructions exist in a patterned way that reflects the structure of unimodal lexical items, it implies that spoken or written language is embedded within a broader holistic multimodal system. Thus, accounting for such phenomena is a requirement for adequately characterizing the structure and architecture of human language. The analysis offered here aims to substantiate our claim that multiple modalities, although displaying modality-specific features, coalesce within a single, holistic cognitive architecture that humans use to construct, integrate and express meaning. Our approach is embedded within the multimodal parallel architecture, a theory proposed in Cohn (2016), and situated within Jackendoff’s (2002) Parallel Architecture for speech. This theory treats different modalities – speech, graphics, and bodily movements – as states emerging from the interaction of parallel distributed structures. Each modality contributes to an integrated conceptual structure using cognitive recourses and combinatorial principles common to all modalities.

2 The Multimodal Parallel Architecture

Following Jackendoff’s (2002) model, the multimodal parallel architecture argues that communicative expressions involve three mutually interfacing components: meaning, modality, and grammar, as depicted in Figure 2.
The multimodal parallel architecture posits modalities as the three basic, equally weighted tools human cognition is equipped with to conceptualize, express and communicate thoughts: graphic structure, phonological (vocal) structure and bodily structure. In addition, the multimodal parallel architecture adds a grammar component to specify each modality’s combinatorial principles—i.e., how a modality packages meaning. These grammars may use modality-specific representations, but they operate with principles of combinatorics that persist across modalities. In each modality, the grammar might range from one-unit sequences, such as single words (vocal), gestures (bodily), or single images (graphic), to fully recursive grammars using categorical roles (parts of speech) and phrasal segmentation. In the spoken or signed modalities, recursive syntactic structures govern the production of sentences, while a recursive grammar at the narrative level structures the visual sequences found in, for example, comics (cf. Cohn, 2013). In addition, other combinatorial principles situated between single unit and recursive structures, such as linear grammars and phrase structure may characterize particular types of sequencing (Jackendoff & Wittenberg 2014), and we elaborate on one such type below.

The multimodal parallel architecture characterizes communicative forms as interactions of one or several modalities, grammar types and conceptual structure. The model’s horizontal dimension allows types of expressions as

---

1 Here, “graphic structure” refers to a representational system governing how lines and shapes go together analogous to phonology, but for the visual-graphic modality. This operates naturally for drawing, but scene-level representations may be mediated by the tools of photographic images. Writing is then the unnatural linking of the vocal modality to the graphic structure to create graphic cross-modal representations of sound.
unimodal triplets of a modality using a type of grammar to organize meaning: verbal languages (phonology, syntax, and conceptual structure), sign languages (bodily structure, syntax, and conceptual structure), and visual languages (graphic structure, narrative structure, and conceptual structure). The model accounts for single unit expressions as forms using only a one-unit grammar, like single words, single images and gestures. The model’s vertical dimension between modalities identify expressions that cross modalities. Written language, for example, employs verbal syntax like speech, but additionally maps phonology to graphics (i.e. orthography, notated in Figure 2 with the dotted arrows between Graphic and Vocal modalities), while beat gestures sync bodily motions to prosodic phonological structures (notated with a dotted arrow between Vocal and Bodily modalities). Multimodality hence follows from combinations of different emergent forms, resulting in multiple activation patterns across the parallel architecture. Hence, while multimodal expressions and language employ analogous means for combining units; what distinguishes multimodality from language concerns the content of what it combines: not only phonology and syntax, but also visual and/or bodily representations.

Interactions within and between components in the parallel architecture can be characterized by breaking apart expressions into their constituent parts and then seeing how those parts are interacting while sharing the conceptual origins (i.e. the ‘growth point’ in McNeill’s (2000) terms). As both meaning and grammar can differ across modalities, multimodal interactions vary depending on what each modality contributes to Conceptual Structure (Cohn, 2016; Bateman, 2014; Martinec & Salway, 2005). Though modalities may convey information in different ways guided by their affordances, within the broader whole of a multimodal interaction, these contributions may be balanced, i.e. modalities share equal “semantic weight” or imbalanced, i.e. one modality’s contribution takes precedence over others. Similarly, expressions may vary in the contributions of their grammars. Co-speech gestures for example are single units combined with full recursive sentences, while visual sequences in comics balance multiple grammars in both text (syntactic structure) and images (narrative structure).

How does this model then account for the possibility of multimodal constructions? The Parallel Architecture defines knowledge of language as a repertoire of constructions: stored associations between pieces of syntax, phonology and meaning at every level of generality/specificity (Jackendoff, 2002; Jackendoff & Audring, 2020). Constructions can be individual words; fixed multiword expressions like idioms (‘let the cat out of the bag’) or stock phrases (‘the benefit of the doubt’) and partially specified constructions (‘what’s that X doing
in det/personal pron Y`). At the other end of the constructional spectrum we find constructional idioms — abstract syntactic patterns consisting entirely of open variables that link to conceptual structures in ways that cannot be predicted by any particular instantiation of its variables. For example, regardless its lexical specification, any instantiation of the resultative construction (NP1-V-NP2-AP) renders an event of causation with the V-slot evoking an instrumental role: NP1 causes NP2 to become AP by V-ing (cf. Goldberg & Jackendoff, 2004). The conceptual structure is therefore linked to the construction as a whole rather than to any of its parts.

Under this view, multimodal constructions qualify as special types of constructional idioms, defined as a \{Modality-slots, Grammar-types, CS\}-triplet. Here, the modality component provides slots for several modality-specific structures and units (graphic, vocal, bodily) instead of only phonology, while the grammar layer contains several modality-specific grammars, instead of only verbal syntax. Any triplet of this kind which uses a patterned schematic form and can be demonstrated to represent meaning independently from its instantiations, qualifies as a multimodal construction; a piece of knowledge that is as such stored in the multimodal lexicon (see also Schilperoord & Cohn, forthcoming).

Prior debates about multimodal constructions have focused mostly on the constructional status of speech-gesture combinations (see for example Cienki, 2017; Ningelgen & Auer, 2017; Ziem, 2017). Opponents argue that gestures often redundantly express verbal meanings. Ziem (2017) for example states that the component modalities should contribute uniquely to a multimodal construction's overall meaning. In addition, it has been stated that the verbal and bodily parts in many speech-gesture combinations co-occur too infrequently to become multimodal constructions.

By comparison, our examples here display co-occurrences of verbal and visual elements that can hardly be considered coincidences or optional. In addition, as highlighted by the Parallel Architecture, meaningful relationships across modalities are just one dimension of multimodal interactions. The constructional nature of a particular multimodal expression should be argued as in Construction Grammar, by demonstrating that the entrenched structural combination of composite parts gives rise to meaning, which is often not overtly expressed (cf. Goldberg & Jackendoff, 2004; Verhagen, 2005). What sets multimodal constructions apart from unimodal constructions is that these structural parts may originate in different coding systems, i.e. modalities. In the next two sections, we will detail the triplet for BA-messages to demonstrate these criteria.
The Nature of BA-Messages

Standard Before-After expressions consist of three components: (1) two images horizontally centered on the canvas; commonly accompanied by (2) the verbal labels ‘before’ and ‘after’; and (3) a slot, usually located below the image dyad, to mention or show the expression’s topic: an advertisement’s product or an event, person or state addressed in a cartoon or meme.\(^2\)

The general meaning of BA-messages is an event of change which is effectuated by a combination of identity and difference. While the two images display some identical object, change is suggested by showing one of this object’s properties, attributes or features in two differing or contrasting qualities or states: the ‘before’ and the ‘after’ state or quality. Figure 1a shows the same lady with wrinkles ‘before’ and without wrinkles ‘after’, so her wrinkles have disappeared. Figure 1b shows Obama with brown hair ‘before’ and with grey hair ‘after’, so the color of his hair has changed from brown into grey.

In addition to change, BA-messages engender an event of causation by implying their topic to be the factor that has caused the change event. Figure 1a hence means “this cream causes your wrinkles to disappear”, while Figure 1b means “having kids’ causes a person to grow older looking” (“a person” in this case indexed by Obama). BA-messages persuade by inviting recipients to attribute valued predicates to the before and after states. Having wrinkles qualifies as ‘bad’, not having them as ‘good’, and so the event of change rendered here qualifies as a change ‘for the good’. And since the advertised product is responsible for this ‘change for the good’, it itself must be ‘good’ as well. By the same token, the states/events addressed by Figures 1b and 1c are claimed to bring about a change from ‘good’ to ‘bad’, and so invite recipients to disqualify their topic.

Putting everything together, the meaning of BA-messages can be specified as a composite of three conceptual relations: a change of state evoked by identity and differences that are visually expressed; a contrast of value valences to be assigned to these states; and a causal relation between the message’s topic and the change-event. This ‘topic causes change for the better/worse’ is claimed to be what every BA-message represents, regardless how the constituting elements are visually or verbally specified. In the upcoming sections we substantiate this claim by detailing (1) the components and relations of the modalities used by BA-expressions; (2) the types of grammar employed by BA-constructions, and how they are balanced; and (3) the conceptual

---

\(^2\) For this analysis a corpus of around 300 advertisements, jokes, memes, and cartoons was sampled, using Google search terms like ‘Before After advertisements/cartoons/memes’.
structure expressed by BA-instances, i.e. the parts of this structure that are compositional and those determined by the construction’s variable slots. In addition, several deviant uses of the construction will be discussed so as to demonstrate that such instantiations invite meaning and persuasion based on the same constructional properties of regular instances. Finally, we argue that the BA-construction is a specified multimodal lexical item within a hierarchy of multimodal constructions that all inherit properties from a more general, fully productive two-image construction.

4 The Before-After Construction: Modality and Grammar

Figure 3 shows the parallel architecture specification of the BA-construction. We begin with detailing the Modality and Grammar components, before moving on to Conceptual Structure.

Following Jackendoff and Audring (2020), we mark correspondences within and between the modality and grammar components with coindices: numbers for the BA-construction’s visual parts and letters for its verbal parts. As we go through the details of Figure 3, these correspondences will be further clarified.

4.1 Modality, Visual

The visual modality tier is specified as the construction’s visual Graphic Structure (GSvi) and its External Compositional Structure (ECS, Cohn, 2020). GSvi is characterized by the two panels (coindices 1 and 2) with dotted circles representing the object and its salient feature or attribute in its ‘before’ and ‘after’ state. A third optional panel (coindex 3) allows visualizing the topic. The specifications ‘W1=W2’ and ‘L1=L2’ express that the two images have equal length and width. Note that the nature of this visual component remains unspecified: it can be drawings, photos, or any other visual media. ECS specifies the layout of the construction as two horizontally aligned panels (1,2) that must be read ‘left-to-right’, and are vertically arranged above the topic panel (3).

4.2 Modality, Verbal

The verbal modality component of the BA-construction is specified as the verbal Graphic Structure (GSve) and Phonological Structure (PS). GSve specifies the orthography, graphemes and fonts used to depict the words ‘before’ and ‘after’, the product name or other pieces of text. In regular instances, similar font is used to depict the words ‘before’ and ‘after’. Graphic depictions of phonological structure concern the words /befor/ and /aftr/, the product name, a slogan or some other verbal expression (coindices a to c).
4.3 Grammar, Visual

The visual grammar component of the BA-construction specifies the construction’s visual morphology (MSvi) and the visual grammar it employs (VG). MSvi specifies the panels (coinicdes 1–3) as so-called monomorphs (abbreviated Mm, cf. Cohn 2018); i.e. visual representations of some isolatable, independent form like the lady’s head with and without wrinkles (cf. Figure 1a). Monomorphs are thus visual analogous to the morphological level of ‘words’ in verbal language at the level of an abstracted form (see Cohn, 2018; Schilperoord & Cohn, forthcoming). The visual grammar used by the BA-construction should be specified within the range of grammars available in the multimodal parallel architecture in Figure 2 – from one-unit grammars to fully recursive grammars. To this end, we adopt Jackendoff and Wittenberg’s (2014) hierarchy of grammars which provides a taxonomy for the basic units and combinatorics a modality can employ to map form and meaning, which has already been applied to visual
sequencing (see Cohn, 2020; Cohn, Engelen & Schilperoord, 2019). Within this hierarchy, the two panels employ a two-unit grammar, since it is constrained to two juxtaposed slots (coindices 1, 2). The visual grammar can hence be formalized as a matrix utterance comprising of two images: \[ \text{utterance} [\text{image}_1, \text{image}_2] \]. These units are ordered by the temporal before-after structure (to be specified in the conceptual structure), while the reading order also matters for the meaning aspect of ‘change for the better/worse’. The third panel does not partake in this two-unit grammar; it is a monomorph that structurally stands on its own as a one-unit grammar.

4.4 Grammar, Verbal

The verbal grammar of the BA-construction is similar to any Parallel Architecture account for verbal expressions. Precisely what is specified depends on whether a specific construction uses words, phrases or made-up sentences. Figure 2a uses only the words ‘before’ and ‘after’, so the MSve specifies these as two words, while the syntactic structure (SS) specifies these as AP (coindices a and b). The phrase ‘the burden of command’ in Figure 1c is specified as a series of words (MSve) forming a complex NP (coindex 3).

The specifications of the visual and verbal modality and grammar tier entail our basic claim: all this structure belongs to the lexical entry of the BA-construction, i.e. the knowledge people have of the construction and use upon encountering one of its instantiations. We now turn to its conceptual structure.

5 The Before-After Construction: Conceptual Structure

This section characterizes the generic ‘caused change’ conceptual structure evoked by the BA-construction and specifies how this structure links to the modality specifications given in Figure 3. We start with the origins of the change event.

5.1 Continuity and Activity

As noted by Hornsby and Egenhofer (2000) ‘In scenarios of change, identity is a key factor in (...) being able to track similarities or differences in objects’ (ibid.: 208). Rather than two lookalikes, the ladies shown in Figure 1a are to be perceived as the same person (i.e. identity) with one feature, her wrinkles, shown in two different qualities: present and absent (i.e. differences). This balance between identity and difference is what provides BA-expressions with the capacity to evoke change, and we shall refer to it as the continuity constraint.
and the *activity constraint* (Cohn, 2020). With only continuity but without activity, the two images would be perceived as identical (one image printed twice); with no continuity, the images would be perceived as depicting different objects. In Figure 1a, the identical lady satisfies the continuity constraint while the presence (before) and absence (after) of wrinkles satisfy the activity constraint. Activity may also concern, for example, a person’s hair (i.e. grey versus colored, bold versus hairy), his or her weight, (i.e. fat versus slim), or physical condition (i.e. out-of-shape versus in-shape), or, in the case of a depicted object like a piece of clothing, with and without stains.

The continuity and activity constraints present a special case of the domain-general *same-except* relation discussed in Culicover and Jackendoff (2012; see also Jackendoff & Audring, 2020). Same-except relations hold between two (or more) objects when these are (perceived as) ‘nearly the same’ but different in one particular salient respect. However, rather than two *type-identical* objects, i.e. “these two ladies are the same, except that one of them has wrinkles while the other one does not”, in the BA-construction continuity refers to *one token-identical* object, i.e. “this is the same lady except that she is shown with and without wrinkles”. So instead of ‘same-except’ judgements, BA expressions invite an ‘identical-except’ judgement, one that will lead recipients to infer that an identical object is shown at two different points in time, and so must have undergone some change.3 (1) gives the full specification of the two constraints.

(1) a. Object $O$ in image 1 is token-identical to $O$ in image 2
b. $O$ has some variable and visualizable property/feature/part/attribute $P$
c. $P$ relates to $O$ in image 1 the same way as it relates to $O$ in image 2
d. $P$ may exist in different, visualizable states $S_1, S_2,\ldots$
e. $S_1$ differs from $S_2$

(1a-c) specify continuity, while (1d-e) specify activity. The activity-part can be further characterized in terms of the kind of property $P$ that has changed,  

---

3 Note that identical-except judgements arise from visual perception only when the compared objects are somehow depicted (drawn, photographed, or else). Two seen objects can never be token-identical. In addition, various presentational aspects invite recipients to compare depicted objects. The two depictions must be sufficiently close to each other and, preferably, be horizontally aligned (cf. Culicover & Jackendoff, 2012: 308). This is completely in line with our specification of GSvi and ECS of the BA-construction. We leave the (empirical) question how far one can deviate from these specifications before identical-except judgements will no longer follow open.
like the size, shape or color of $O$. Furthermore, the way $S_1$ and $S_2$ differ from each other determines the nature of change. Here we can distinguish addition, elimination, and alteration (cf. Culicover & Jackendoff, 2012, 309; Schilperoord, 2018). An addition happens when $S_2$ is a sort of ‘extra’, like for example in a hair restoration advertisement that shows hair on a ‘before’ bald head. An elimination occurs when $S_2$ concerns something that has been removed. It creates contrasts, like the presence vs. absence of wrinkles on the lady’s face in Figure 1a. An alteration occurs when $S_1$ and $S_2$ concern a state of $P$ that has altered in some respect, like for example Obama's grey versus black hair in Figure 1c.$^4$

5.2 Caused Change

Next to evoking an event of change-of-state of an object, the BA-construction also links this event to its topic. In advertisements the construction invites conceptualizing the recommended product or service as the causal factor that brings about change, while cartoons suggest some socio-political actor or event as the factor that has caused change. However, whereas the change-event is evoked by the two images (provided these honor the continuity and activity constraints), the cause-event is not overtly expressed and should thus be considered entirely constructional. The conceptual structure of the BA-construction can thus be formalized as (2), where underlines set the constructional part apart from what is overtly expressed.

$$\begin{equation}
(2) \quad \text{CS} \ [\text{event change \ldots}] \end{equation}$$

The [event change \ldots]-part of (2) can be further formalized as realizing the conceptual function [go ([object]), \ldots] and the change itself as a path: an object ‘goes’ from a certain state to another state with regard to a single property, feature or attribute (cf. Jackendoff, 1990: 26, 96ff, 112). This is expressed in (3), where $O$ denotes the object and $S_b$ and $S_a$ denote the before and after states.$^5$

$^4$ There will probably be other ways of change, but we leave this question open. Another empirical question is how many $P$’s can be subjected to change at once before identical-except judgements no longer arise. We suspect not many, which is nicely testified by parodying usages of the construction. A BA-parody is an instantiation in which ‘too many’ $P$’s have been subjected to change, like a fat man ‘before’ and young Arnold Schwarzenegger ‘after’.

$^5$ Rather than formalizing CS as $X$ cause $Y$ to become $Z$, like proposed for the resultative construction, we have CS express a path-function from \ldots to \ldots so as to account for the change-aspect of meaning and the fact that the two images in de BA-construction visually express both states. Note that under the alternative characterization, the CS in Figure 1a would be product cause woman to become unwrinkled, which in our view is not what the message expresses. This particular CS would be the proper one had the advertisement only showed the ‘after’-state image of the lady.
The index $\text{Prop}$ attached to the change-function $\text{GO}$ expresses that it is one particular property or attribute of entity $O$ that is the object of change. By applying (2) and (3) to Figures 1a and 1c, we can express their meaning as (4a) and (4b).

(4) a. $\text{CAUSE (}[\text{entity PRODUCT}] \text{GO}_{\text{Prop}} (\text{[entity WOMAN]},$

$b. \text{CAUSE (}[\text{state BURDEN-OF-COMMAND}] \text{GO}_{\text{Prop}} (\text{[entity OBAMA]},$

These conceptual structures adequately express the generalization that the causal factor is always the topic of the message and can be an entity (4a) or a state (4b). In the Conceptual Structure of Figure 3 we have linked the various parts of CS to the specifications of the modality and grammar tier of the BA-construction. Numeral coindices again express visual-conceptual correspondences and letter coindices express verbal-conceptual correspondences. Coindices 1 and 2 link the depicted object, and its property rendered salient, i.e. monomorphs 1 and 2 (MSvi), and the two images that make up the utterance (VG), to $[\text{entity — }]_{1,2}$ and the two states $[\text{state — }]_{1}$ and $[\text{state — }]_{2}$ at CS. Coindices a and b represent MSve/SS/CS interfaces: $[\text{Wd — }]_{a} / [\text{AP — }]_{a} / [\text{state — }]_{a}$ and $[\text{Wd — }]_{b} / [\text{AP — }]_{b} / [\text{state — }]_{b}$ respectively. Coindex 3 links the monomorph used to depict the topic to the $[\text{entity TOPIC}]$-part at CS, while coindex c does so for the word-slot and syntax used to signify the topic verbally. Coindex 4 associates the VG-specification $[\text{utterance [image], [image]]}$ and the CS-part $[\text{GO}_{\text{Prop}} (\ldots \text{FROM (__) TO (__)})]$. This motivates the use of a two-unit ordered grammar as it links the temporal ordering of the images to the FROM- and TO-state. Finally, the $[\text{CAUSE (\ldots)}]$-part of CS is not linked to any element of the construction’s modality and grammar tiers since $[\text{CAUSE (\ldots)}]$ isn’t overtly expressed but the constructional effect of the image/text unions and of the conditions imposed upon the kind of images as expressed by the continuity and activity constraints. This accounts for $[\text{CAUSE (\ldots)}]$ being expressed by any instance of the construction.

6 There is an obvious parallel here with, for example, the way the verbal resultative construction (NP V NP AP) expresses causation.
5.3 Value and Valence Assignment

A full account of the conceptual structure of the BA-construction calls for specifying how, besides evoking a process of change, the construction construes this change as one ‘for the better’ or one ‘for the worse’; i.e. how the construction invites persuasion. This evaluative aspect of meaning operates on attributing a particular type of value to the before and after states with contrasting valences. As noted before, the advertisement in Figure 1a employs the intersubjective notion that having wrinkles looks or feels ‘bad’, whereas a smooth skin looks and feels ‘good’. So, rather than a matter of mere appearance, the states shown in the images critically suggest some negative affect to have transformed into a positive one. And since the product is claimed to be capable of bringing about this change from negative to positive, it must itself be ‘positive’; hence worth the purchase.

To account for value attribution in conceptual terms, we draw on Jackendoff’s (2007) definition of values as ‘(...) a conceptualized abstract property attributed to (conceptualized) objects, persons or actions (...)’ (Jackendoff 2007: 376, our italics). Instead of being inherent properties, values are conceptual categories which are attributed to persons, objects or events ‘by us’, based on some observed facilitative circumstance. Value attribution can thus be modelled as an input rule: “if such-and-such is the case with X, or if X does such-and-such, then attribute to X such-and-such a value”. Depending on the input rule’s antecedence, various types of values can be distinguished, like for example quality, prowess, utility or affect values. When an advertisement claims its product to be exceptionally good in what it does (i.e. its proper function), it invites attributing quality value; when the product is claimed to yield a benefit, its utility value is claimed, and when it provides pleasure, affect values are at stake. By the same token, a cartoon that mocks a politician invites attributing (negative) prowess or normative values to their topic. These examples furthermore indicate that values come with a

---

7 Such valence assignment, i.e. that the face crème advertisement implies a change for the ‘good’, no doubt also interacts with ‘frame-semantic’ knowledge, i.e. the cultural frame of facial beauty. Such frames are assumed as part of Conceptual Semantics, and compatible with our analysis of value assignments. Values are assigned based on certain input rules which may implement such cultural frames. We assume all examples discussed in this paper are understood with reference to encyclopaedic knowledge, i.e. that a face without wrinkles qualifies as ‘good’ and ‘beautiful’ or that severe headaches are metaphorically understood as being hit with a hammer on the head. However, we do not attempt to render explicit how contextual knowledge interacts with our expressions, but instead to identify the constructional properties of the BA-construction with which contextual knowledge can interact.
valence: positive or negative. Having a smooth skin has a positive affect-value, whereas having wrinkles has negative affect-value.\(^8\)

Because the two images in BA-messages depict a change of state which attract value-attribution with opposite valences, and because the construction relates the topic causally to the depicted change of state, the images function as the antecedence of the input rule for attributing value to the topic: if the topic causes the valence of value X to change from negative into positive, then attribute to it value X/Y... also with positive valence. The type of value attributed to the topic may differ from those suggested by the images. While the images in Figure 1a attract affect value, the product is claimed to have quality value – it is ‘good’ in what it does. In fact, most advertisements in our corpus invite attributing quality value to the recommended product or service, regardless of the values suggested by the images.

The formalities of value attribution and its links to the construction appear in (5), copying only the CS specification from Figure 3. Value attribution is marked provisionally as VA. Small capitals denote value types, i.e. Quality, Utility, Prowess, Affect,..., while symbols + and – denote positive and negative valence.

\[(5)\]

\[
\text{CS} \left[\text{cause} \left(\left[\text{entity topic}\right]_{3,c} \left[\text{goProp} \left(\left[\text{entity \ldots}\right]_{1,2} \left[\text{from} \left(\left[\text{state \ldots}\right]_{1,b}\right)\right]\right)\right]\right)_{4,5}\right]_{5} \\
\text{VA} \text{ if } \{Q/U/P/A... \text{ [before]}_{1,a} = - & Q/U/P/A... \text{ [after]}_{2,b} = + \} \\
\text{then } Q/U/P/A... \text{ [topic]}_{3,c} = +\]

Note that the causality suggested by the If-Then input rule at VA is motivated by the entire CS [cause (...)], which explains why coindex 5 links VA to the entire CS. VA represents a necessary addition to the BA-construction’s CS as its communicative intentions reach past mere description and always include evaluating the message’s topic. Even a recipient who manages to construe the ‘caused-change’ meaning of Figure 1a, cannot be said to have fully grasped the intended message of the ad when the evaluation is not included in his/her interpretation. Therefore, the Then-part of the input rule in (5) marks the conceptual ‘endpoint’ of the message as intended.

Valences specified in (5) will be a matter of the pragmatics of the genre. Unlike advertisements, the end point in many editorial cartoons employing the BA-construction concerns a value with negative valence rather than a positive one. The cartoon in Figure 1c for example invites attributing utility value with negative valence to its topic ‘the burden of command’: ‘if this is what the

---

\(^8\) For a complete overview, see Jackendoff (2007: 280–303).
burden of command does to Obama, it is ‘bad’ for him’. Hence, in this case the valences of the ‘before’ and ‘after states are reversed: from positive-before to negative-after. However, generalized across genres, the valence of the value attributed to the topic will be similar to the one attributed to the ‘after’ state. If the valence attributed to the after-state is negative, so will be the valence attributed to the topic.

6 Deviant Cases

The aim of this section is to provide additional substance to our claim that BA-expressions instantiate a multimodal lexical/constructional item. We do so by discussing several ‘deviant’ cases; i.e. cases that deviate from one or more of the prototypical BA-construction’s properties. We demonstrate that despite those violations, meaning-making still critically involves the BA-construction’s standard properties. The common thread running through this section will thus be that recipients do not process and interpret cases like these ‘on their own’, but as meaningful deviations from the standard BA-construction as outlined above.

Before we start, a brief remark on notation. To highlight their significance, several analyses offered below need some way of formalization. For brevity, we shall use the simplified notation in Figure 4a instead of the full formalization presented in (4) to (5).

This notation skips the modality and grammar structures of the construction, and captures only its conceptual and valence structure. The three main parts of the construction are given between square brackets. The topic of the message appears as ‘entity 1’, with a certain value \(v\) assigned to it with positive or negative valence. The two images appear as ‘entity 2’ showing it in the before and after states \(x\) and \(y\). Both states are assigned a (similar) value \(v\) either with negative (–) or positive (+) valence. Change is expressed by the doubly pointed horizontal arrow while Cause is expressed by the vertical arrow.

6.1 Violating Modality or Grammar Specifications

Consider the advertisements in Figure 5a to f, which illustrate violations of the BA-construction’s modality or grammar specifications.

Figure 5a, an ad promoting non-surgical fat removal, violates verbal syntax specifications (SS in Figure 3) by using the adverbs ‘before’ and ‘after’ as predicate nouns (“You have been a before long enough, become an after”) flavoring the terms with a generic stance. To properly construe this punning, awareness of
the BA-construction and its typical adverbs is needed (imagine what a recipient ignorant of the BA-construction would make of this). The advertisement also violates the visual Graphic Structure (GSvi) specification that the two images/panels have equal length and width – the before-panel is considerably wider than the after-panel. Given the topic, this deviation clearly makes sense, but one that can only be appreciated with knowledge of the standard constructional features. The same kind of violation is present in Figure 5b, also promoting solutions to obesity.

Although similar fonts of the words ‘before’ and ‘after’ is specified at the verbal GS, Figure 5c illustrates how using different fonts transfers these words to
Figure 5  Advertisements which deviate from the canonical Before-After-construction. They manipulate dimensions of a) panel size and syntax, b) panel size, c) font, d) spelling, e) external compositional structure (layout), and f) a lack of continuity.
the visual structure of a BA-expression – the imprints “before” and “after” iconically signify the states of being overweight and being slim. Again, it is hard to see how this deviation makes meaning without reference to the standard specification of the BA-construction.

Figure 5d promotes a literacy program and violates the convention that words used in public messages should be spelled correctly. Incorrect spelling of the words ‘before’ and ‘after’, i.e. violating the PS-GSve mapping for orthography, is used here to signify the state of illiteracy and the state of being able to spell words correctly – again a deviation with obvious semantic intent. Figure 5e promotes the services of an eye surgeon clinic, and violates the external compositional structure specification at the visual modality tier (/ Horizontal Panel1 – Panel2/) by replacing it with the Snellen chart to evoke the before-after states. Again, this makes perfect sense, but only to recipients capable of monitoring it as deviation – i.e. who have knowledge of the standard construction.

6.2 Violating the Continuity or Activity Constraint
A quite spectacular deviation concerns BA-messages containing images that violate the continuity or the activity constraint. Consider Figure 5f, an advertisement that promotes an esthetic dentist. In this advertisement, the navigational left-to-right order of the ‘before’ and ‘after’ panels is flipped. However, since all basic constructional properties remain intact, the order is consistent with a script that reads right-to-left, implying it can follow that of a culture’s writing system (see also Cohn, 2020). The more creative deviation comes from the content of the panels though. Instead of showing a person’s teeth before and after treatment, the message shows a split-screen of two sides of an apple, each with a particular biting contour. Although the image does suggest a single (i.e. token-identical) apple, the before and after states depict two different biting contours, not one that has changed relative to some continuous object over time. To be able to make sense of this message as a BA-message, hence to understand what has changed, the critical insight is that change applies here to a non-shown object. The bite marks are metonymic of the teeth, and their properties, that have bit the apple (see the doubly pointed vertical arrows in Figure 4b). Both the object’s relevant property (some client’s teeth) and its before-and after-state (having ‘bad’ and having ‘good’ teeth) have to be inferred from what is actually visualized and, of course, from the message’s topic. Again, however, this inference is driven by the standard BA-construction’s caused change meaning: S1 and S2 concern altered states of a person’s teeth which are caused by the advertised dentist. Once this inference is properly established, it can be deduced that the images function as an indexical sign of what has
actually changed: the irregular contours of the ‘bite’ in the apple on the right index ‘bad teeth’ while the regular contour on the left index ‘good teeth’.9

Figure 4b accounts for this way of restoring the continuity and activity constraints. The inferred level of meaning is marked by a dotted frame. The BA-construction anticipates the inference in general terms (i.e. \([\text{entity } \langle x, v (-) \rangle \rightarrow \langle y, v (+) \rangle]\)), but details must be inferred from the top layer. The biting contours, which relate in terms of contrast, allow inferencing an irregular as opposed to a regular configuration of teeth, and so satisfy the activity constraint. This also yields the indexical relation linking the inferred change of state and the images-as-shown: the configuration of teeth is indexed from the biting contours in the apple. And since the teeth satisfy the continuity constraint, the evoked change of state is accounted for with its associated value(s) and contrasting valences. The inferred conceptual structure feeds the antecedent of the input rule for value attribution. Negative affect and utility values are attributed to the before-state – it ‘feels’ and ‘is’ bad for one to have teeth that produce biting contours like this – while the after state attracts the same values but with positive valence. This valence is carried over to the topic which yields a positive prowess value to the advertised dentist as conceptual end-point.

This inferential pattern occurs quite frequently in our corpus. In all these cases the depicted objects do not apply directly to the inferred before-after pattern (hence violating the continuity and activity constraints), but are indices, allowances, byproducts, or metaphors of the change that occurs to an unseen object. Consider the examples in Figure 6a to f.

Figure 6a promotes an aspirin and shows two hammers, one with an iron head (labeled ‘before’) and one with a soft ‘pillow’-like head (labeled ‘after’). Instead of one token-identical hammer that has undergone change, we see two hammers differing in the material of the head. The to-be-inferred before-after change is again suggested by the topic: aspirins remedy headaches, so the relevant state is ‘headache’ which has been changed from ‘present’ to ‘absent’. In this case the relation between the inferred change-structure and the actual images is one of analogy. The two hammers metaphorically visualize how bad it feels having a headache; i.e. like (being hit with) a hammer with an iron

---

9 Depending on one’s preferred theoretical framework, the inferred level of structure and the one representing what the two images actually depict can also be modelled as a blend between structures from different domains: one derived from the general specification of the BA-construction, and one from bottom-up processing (cf. Fauconnier & Turner, 2002). In addition, the bite marks in the two apples can be said to be metonymically related to the intended properties of the teeth. However, we make no claims about generalized types of construal since our focus here is on formalizing the specific predicates that are involved.
head, and how good it feels once freed from it: like a hammer with the soft head. Figure 4c is the full conceptual structure.

Figure 6b asks people to donate money for the blind. The two images represent a contrast by depicting only the words ‘before’ and ‘after’: before appears in braille (left) and after in the graphic alphabetic script (right). Altering states that are to be inferred concern someone forced to read braille and the same person being able to see again. The relation between the inferred change-structure and the images is thus one of allowance.

**Figure 6** Example Before-After constructions that manipulate the inference of the canonical pattern. These examples evoke a) inferred metaphorical states, b) inferred allowance, c) inferred causality, d) violation of continuity, e-f) puns.
As another example, Figure 6c requires inferring a change of states from bad to good behavior of a dog, a change caused by the recommended ‘dog whisperer’. This BA-instance violates continuity since we see two different shoes. The inference anticipated is of causality between the depicted images: dogs behaving badly ruin shoes while good dogs leave them untouched.

The cleverly designed advertisement shown in Figure 6d exemplifies a violation of the activity constraint. This message repeats the before state as after state, in effect showing the same image twice. The violation makes sense given that the advertisement claims the sustainable quality of the advertised marker’s ink. The images present a visual pun in that they mimic a typical BA-advertisement for laundry detergent – a piece of stained clothing before and shining clean after – except that the stain is still there. The absence of activity engenders the messages that no matter what laundry detergent one uses, the ink on the shirt produced by the advertised product will never vanish, thus suggesting its outstanding quality value. What has changed here is that the expected change hasn’t occurred – again the kind of interpretation that is driven by the standard BA-construction.

6.3 Puns

Figure 6d hints at the possibility of BA-puns. Allowing for punning indeed testifies to an expressive pattern’s constructional nature (cf. Giora et al., 2004; Arts & Schilperoord, 2016). BA-puns instantiate the standard construction, including its original effect, but add a novel layer of meaning that both employs this effect and at the same time mocks or denigrates it. Figures 6e and 6f are examples. Figure 6e intends to raise people’s awareness of the consequences of suffering from Multiple Sclerosis. Like Figure 6d, the images evoke a laundry detergent advertisement except that the ‘filthy’ piece of clothing is displayed as the after-state. Hence, we get a ‘change for the bad’, visualizing what may happen when one suffers from Multiple Sclerosis. Figure 6f promotes a Spanish cable channel (‘Channel 13’) known for its terror, mystery and action TV series and movies. The images show the typical male used in shampoo advertisements claiming it brings back your brown hair, but again appear in an unanticipated order: instead of appearing as before-image, i.e. the one that attracts negative valences (having grey hair feels bad), it is shown as the after-image. Apparently, the broadcasting is so good that its sheer suspense will have turned your brown hair into grey after a year of watching it. The grey-hair state is hence used for the exact opposite purpose: to attract positive valence. Getting this message again crucially involves reference to the regular constructional properties of the BA-construction.
Hierarchy of Juxtaposed Image-Constructions

Having argued for the BA-construction as a multimodal lexical entry that people activate upon encountering instantiations, we next question how this entry is mentally represented. We argue that instead of an entry-in-isolation, the BA-construction’s formal properties suggest it to belong to a broader class of ‘two-juxtaposed-image’-constructions which is organized as an inheritance hierarchy of the sort expounded by Diesel and Tomasello (2000) and Verhagen (2005: 11ff) among others. Consider Figure 7.

The region marked by the dotted frame captures the kinds of BA-patterns that have been so far discussed; ranging from the productive schematic pattern ‘above’ to fixed instances of usage ‘below’. Repeated exposure to examples like Figure 1a, leads gradually to abstracting away from specific usage instances to lexicalizing the abstract construction. Knowledge of this entry comes to exist independently from actual instances of usage. In addition, repeated experience

**Figure 7** An inheritance hierarchy for juxtaposed image constructions.
with the kinds of deviant instances of the construction discussed in the former section may lead to storage of specialized, yet still abstract lexical items, like the inferred before-after layer (see the intermediate levels in the dotted frame in Figures 4b and 4c). Because those more specified items inherit the basic BA-constructional properties, they contribute to a further entrenchment of the canonical construction.

While less entrenched BA-instantiations may require analytical processing, conscious attention and inferencing (e.g. Figures 6a to f) to find out why and to what effect the regular constraints are not satisfied, at the bottom of the network we find BA-expressions so prevalent that they are stored as multimodal ‘stock phrases’, analogous to verbal idioms or larger-than-word fixed expressions like ‘to kick the bucket’, or ‘the benefit of the doubt’ (Schilperoord & Cozijn, 2010). Although these cases all instantiate the general BA-construction, processing and understanding is likely to happen ‘en bloc’, that is, without the need to pay attention to the parts. Examples include highly frequent before-after advertisements, like those promoting laundry detergents, hair restorers, face creams, gyms, fat removal services and the like. We also find them among Internet memes instantiating the BA-construction based on specific contexts, such as the countless ‘me voting in 2016, vs. me voting in 2018’ before-after meme which spread across social media in the days leading up to the 2018 United States midterm elections, only to vanish after election day (itself seemingly a variant of a ‘me T1 vs. me T2’ meme). However, although repeated exposure to canonical examples may lead to storing more abstract patterns, they don’t disappear from the lexicon (cf. Jackendoff & Audring, 2020: 61ff).

Next, consider what might be present ‘above’ the dotted frame in Figure 7. The canonical BA-construction branches from an even more abstract two-juxtaposed images plus topic slot construction; a construction we refer to as the Two-Image, or 2I-construction. The 2I-construction is shown as Figure 4d, where the capitol E stands for image of Entity, R for Relation, T for Topic and (V ±) for Value type + valence.

The 2I-construction abstracts away from the BA-construction by offering variable (instead of fixed) slots for various types of conceptual relations, types of images, values and valence configurations. The continuity and activity constraints no longer apply, and the E1 and E2 may be linked by relations other than ‘change’ (see the open slot R1). R1 can express relations of contrast, analogy, causality, or identity. In addition, the variable slot R2 may also express relations other than Cause, while it allows T to relate to the entire complex {E1, E2}, like in the BA-construction, but also only to E1 or E2. Third, the slot (V ±), i.e. the type of value the expression invites assigning to its components, is free: normative, affect, utility, quality, prowess, or other, and so is the nature
of the valences: positive or negative. Finally, the variable slots $E_1$ and $E_2$ of the 2I-construction allow the insertion of almost any conceivable type of image, including pairs that do not satisfy the continuity or activity constraints – i.e. the defining property of the BA-construction.

The fully productive nature of the 2I-construction comes with a processing cost. While interpreting BA-messages is guided by the much more constrained properties of the BA-construction, expressions instantiating the 2I-construction have many degrees of interpretative freedom. To reach full understanding, the reader has to establish the visual/semantic details of the aligned images ($E_1$ and $E_2$), find out how they relate to each other, and to the topic T ($R_1$ and $R_2$), and how the various slots for value assignment should be fixed.
Let us consider some of the expressive potential of the 2I-construction, and see how further ‘outward’ the hierarchy extends. Figures 8a to f provide examples. Figure 8a illustrates contrast. The heap of feathers and the bucket of tar metaphorically visualize a car’s ‘sticking’-quality without (valence –) and with (valence +) the recommended tire, with quality as the suggested value type. The topic hence relates to the right-side image as analogy; “our tires stick to the road like tar”. Figure 8b illustrates analogy between E1 and E2. Here, the topic (a guitar) is identical to one of the images (T = E2) to produce “our guitar energizes like an atomic bomb”, hence the valence of the affect value suggested by the images is positive. This case also nicely illustrates the interactive nature of metaphoric relations between a source (the bomb) and target (the product). The target gets the mapped property of the source but at the same time selects the relevant property of the source. This explains the invited affect value with positive valence. When the bomb image is used in isolation, this positive valence would be unlikely.

Figure 8c, a dandruff protection shampoo advertisement, has images which relate in terms of alternative means to establish a certain goal (cf. ‘choose’) with differing quality values. Like in Figure 8b, the topic here is identical to E2. The topic in Figure 8d (a fast food chain, but that is not the advertiser) links the images in terms of pretense versus reality: E1 shows the kind of hamburger the food chain pretends to deliver (affect, +), while E2 shows us the hamburger we actually get (affect, – at least according to the advertiser). The message hence confronts contradictory perspectives E1 and E2 which explains the sense of ‘discovery’ engendered by this message. Figure 8e advertises a car brand (“the best or nothing”) by showing Albert Einstein in a serious ‘logical’ mood (E1, prowess +) and in a playful creative (E2, affect +) mood, labeled as corresponding to the left and right brain hemispheres. Here we need a complementary relation to link the images: Einstein was a unique human being because his mind united these traits. To connect the images to the topic the recipient must again infer a covert layer, which represents the two qualities of the recommended car with similar complementary relation. T relates to these covert layers in terms of possession, while the possessed entities relate to the images in terms of similarity. Finally, Figure 8f advertises a brand of wine (‘some things get better with time’), linking E1 and E2 in terms of change, just like the BA-construction. However, rather than as the causal factor that brought about the change, the product is itself the object of this process of change-for-the-better. The images show the

---

10 One could argue that this is a Before-After case as well, like in Figure 6d, with in inferred layer signifying the before state ‘no grip’ and the after state ‘firm grip’, brought about by the product. We leave the issue open.
famous Hiroshima atomic bomb photograph next to one of a mushroom. The suggested change does not concern these objects but the denotation of the word ‘mushroom’: then the word referred to an atomic bomb explosion, now to a mushroom proper. The ‘then’-product state and the ‘now’-state relate to these changed denotations in terms of similarity: bad once, good now.

In terms of Jackendoff and Audring (2020: 107ff) the relation between these zl-types can be captured as ‘sister’ relations. They generalize to a more abstract level of ‘sister schemas’ (see the upper layers of the inheritance hierarchy in Figure 7). By way of rounding of this section, we shall attend briefly to what may be ‘sister schemas’ in the inheritance hierarchy. Consider the multimodal message in Figure 9a.

Figure 9a was part of a 2005 advertising campaign for the Chicago Tribune newspaper with many variant advertisements employing the same three-panel template shown here. All ads highlighted the benefits of sections of the newspaper, with the header “Chicago Tribune, what’s in it for you?”. In an analysis of these advertisements, Cohn (2010) notes that they all follow the same pattern shown in Figure 9a. The first panel sets up some situation, like a person descending an escalator. The second panel features this person reading a section of the newspaper or website labeled in text (here: ‘career builder’). The final panel features an image nearly identical to the first panel, though altered in some way: the same person now ascends the escalator. The strip employs a narrative structure: it starts by setting up an event (the ‘initial’; panel 1), followed by a causative force in the climax (the ‘peak’; panel 2), and ends by showing what has changed (the ‘release’; panel 3). Cohn’s analysis of the grammatical and conceptual structure of the strip is (6), where coindices mark grammarmeaning correspondences.

(6) \[
\text{VG} \quad \text{arc initial}_1 - \text{peak}_2 - \text{release}_3 \\
\text{CS} \quad \text{EVENT} \quad \text{CAUSE} \left( \text{EVENT} \text{READ} \left( \text{PERSON}^a, \text{CHICAGO TRIBUNE} \right)_2^1 \text{INCH} \left( \text{Y}(\alpha)_{1,3} \right) \right)
\]

The formalism for CS in (6) employs the function \text{INCH}(ative) to describe an event in which some entity (\text{PERSON}) undergoes some change (Y) to end up in some altered state. The event of the person reading the newspaper, i.e. the climactic ‘peak’ panel, is analyzed conceptually as the causal force that inputs the \text{INCH}-function. The canonical narrative schema of initial - peak - release sequence in these instances can therefore be said to facilitate the basic conceptual functions \text{CAUSE} and \text{CHANGE} of the BA-construction. The initial and release panels depict the \text{FROM} [...] TO [...] function while the peak panel evokes the conceptual structure \[\text{CAUSE} \left( \text{[TOPIC]}, \text{GO} \left( \text{[entity ...]} \right) \right)\],
with the newspaper being the topic and the entity being the person reading a particular newspaper’s section. In addition, neither the BA-construction nor the narrative pattern underlying Figure 9a overtly express the causal function. These conceptual similarities are further testified by the fact that the Chicago Tribune strip translates straightforwardly into a BA-message, as

![Figure 9](image_url)

**Figure 9** An advertisement for the Chicago Tribune Newspaper using a) a causative sequence with a basic narrative structure and b) a manipulated version converting it to a BA-construction. These create c) two of several options for conveying causative sequences.
in the manipulated version in Figure 9b, instantiating the specialized covert layer BA-item, with the peak panel occupying the topic-slot, and the initial and release panels occupying the before and after slots. Indeed, reverse translations are also possible: Imagine an image inserted between the before and after images of Figure 1a showing the lady actually using the recommended product. This sequence would translate into the same initial - peak - release sequence underlying the Chicago Tribune strips. The covert layer the reader must infer follows from the metaphoric nature of the images: the person goes down to the left before, and up to the right after, which metaphorically expresses the two career states the ad refers to. All this suggests that the semantic functions of the general caused-change conceptual structure can be performed by several 'dedicated' multimodal constructions (see Figure 9c).11

An interesting hypothesis, albeit one we leave for further research, is that instances of these constructions all evoke caused-change and can thus be mutually ‘translated’ by adjusting conceptual parts to the ways these construction map meaning to their graphic structure and grammar: BA: juxtaposed images + 2 unit ordered grammar, IPR: panel sequence + recursive grammar.

8 Conclusion

In this paper, we have made the case for BA-expressions to constitute a multimodal construction which as such is part of the multimodal mental lexicon. While the construction offers the possibility of construing meaning build around the caused-change conceptual structure, it also belongs to an inheritance hierarchy of less specified, more productive 2I-constructions. As we have demonstrated, the intended meaning of ‘novel expressions’ invoking this general 2I-construction should be discovered by making use of the visual/semantic details of the aligned images, the context, i.e. the topic of the 2I-message in question and the construction’s various relational possibilities. Finally, we demonstrated that the BA-construction intersects with a hierarchy of caused-change constructions. These patterns and their relations provide

11 A third visual construction that is capable of suggesting caused-change is discussed in Schilperoord and Van Weelden (2018), and operates on incongruent shadow depiction. The ad in question is one for an energy supplement and shows a young, energetically running woman dressed in sports outfit. She casts a long ‘impossible’ shadow, projected behind her on a wall, which captures her while taking a break, and bending forwards with her hands on her knees. Caster and shadow underpin the idea of one protagonist and apparently express before-after states, the latter caused again by the powerbar.
further support that linguistic patterns arise beyond the verbal domain alone. Such evidence requires that linguistic theories be capable of accounting for multimodal expressions representing the full range of different modalities’ expressions.

References


Schilperoord, Joost and Neil Cohn. Forthcoming. Let there be ... visual optimal innovations. Making visual meaning through Michelangelo’s Creation of Adam. Visual Communication. (Accepted for publication).


Ziem, Alexander. 2017. Do we really need a multimodal construction grammar. Linguistics Vanguard, 3(s1), 1–9.