JOURNAL OF COGNITIVE PSYCHOLOGY, 2016
VOL. 28, NO. 5, 559–574
http://dx.doi.org/10.1080/20445911.2016.1179314

Meaning above the head: combinatorial constraints on the visual vocabulary of comics

Neil Cohn[a], Beena Murthy[a] and Tom Foulsham[b]

[a]Center for Research in Language, University of California, San Diego, CA, USA; [b]Department of Psychology, University of Essex, Colchester, Essex, UK

ABSTRACT
"Upfixes" are "visual morphemes" originating in comics where an element floats above a character's head (ex. lightbulbs or gears). We posited that, similar to constructional lexical schemas in language, upfixes use an abstract schema stored in memory, which constrains upfixes to locations above the head and requires them to "agree" with their accompanying facial expressions. We asked participants to rate and interpret both conventional and unconventional upfixes that either matched or mismatched their facial expression (Experiment 1) and/or were placed either above or beside the head (Experiment 2). Interpretations and ratings of conventionality and face–upfix matching (Experiment 1) along with overall comprehensibility (Experiment 2) suggested that both constraints operated on upfix understanding. Because these constraints modulated both conventional and unconventional upfixes, these findings support that an abstract schema stored in long-term memory allows for generalisations beyond memorised individual items.

ARTICLE HISTORY
Received 8 October 2015
Accepted 12 April 2016

KEYWORDS
Visual language; visual morphology; visual metaphor; emotion; comics

1. Introduction

Visual culture has long drawn from the "visual vocabulary" of comics' unique graphic representations. For example, iconic lightbulbs floating above the head no longer represent a source of light, but convey the meaning of inspiration. Meanwhile stars, a conventional symbol representing a celestial object, mean dizziness when above someone's head. Also, hearts or dollar signs may substitute for someone's eyes—a form now pervasive in "emoji" used in digital text-based communication. These elements have frequently been compared to lexical items in language (Cohn, 2013; Forceville, 2011; McCloud, 1993; Walker, 1980), with competing theories echoing similar debates as in psycholinguistics. "Visual morphemes" (Cohn, 2013) like these have generally been viewed as unique and individualised representations (Kennedy, 1982; McCloud, 1993; Walker, 1980), possibly with metaphoric or embodied meanings (Forceville, 2005, 2011; Kennedy, 1982; Slepian, Weisbuch, Rutchick, Newman, & Ambady, 2010). However, recent work has argued that many of these graphic signs may belong to classes of abstract schema stored in memory, beyond item-based instances, and use combinatorial structure beyond simple concatenation (Cohn, 2013). Here, we explore this hypothesis specifically for the particular class of "above the head" meanings.

Some work has recognised that the context and position of visual morphemes matters for their interpretation (Cohn, 2007; Forceville, 2011; McCloud, 1993). For example, McCloud (1993) noted that curvy lines above coffee indicate heat, but curvy lines above trash indicate a bad smell. Similarly, Forceville (2011) noticed that a spiralling "twirl" above a character's head meant dizziness, but twirls next to a body showed motion. Stars also vary in meaning: when above the head they mean dizziness, but in the eyes they indicated a desire for fame (Cohn, 2007, 2013). Thus, context matters for how these morphemes are interpreted. Because of this context sensitivity, it has been hypothesised that comic reading experience is necessary to understand these signs (Forceville, 2011). Although, some research has suggested that comprehension of these representations is modulated by age and experience reading comics

CONTACT Neil Cohn neilcohn@visuallanguagelab.com

(Nakazawa, 1998, 2004, 2005; Newton, 1985), this hypothesis has yet to be systematically tested.

In recent work, we have theorised that these form-meaning pairs are stored in memory analogously to lexical items in a language (Cohn, 2013). To create meaning with these "morphemes", this "visual language" uses similar combinatorial strategies as in the morphology of verbal languages: speech balloons attach one sign to another (affixation), eyes become hearts or dollar signs to replace one sign with another (substitution/suppletion), and multiple body parts repeat elements to show movement (reduplication). It is important to stress that this comparison between the visual language of graphics and verbal language of speech is an analogy of function only. Speech balloons are not meant as an affix in exactly the same way that *-ness* serves as affix in the word *awareness*. Rather, the analogy here is that speech balloons (and others) are "bound morphemes" that must attach to another "root" morpheme, such as a "speaking person" (Cohn, 2013; Engelhardt, 2002; Forceville, 2011). While the roots may exist independently without the affix, bound morphemes must attach to another morpheme—they cannot stand alone. This implies that a similar strategy of "attachment" governs the combinatorial structure of both verbal and visual domains, but whether or not they involve a common underlying cognitive process is an open question. Recent work has shown that similar neurocognitive responses are evoked by violations to the "grammar" of sequential images and syntax in sentences (Cohn, Jackendoff, Holcomb, & Kuperberg, 2014; Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012; Sitnikova, Holcomb, & Kuperberg, 2008), and thus it is at least conceivable that similar operations might govern combinatoriality on units within those sequences (e.g. Cohn & Maher, 2015).

"Above the head" meanings provide particularly rich examples of visual affixation. These elements have been named "upfixes" because they are graphic affixes that go "up" from a head (Cohn, 2013). As depicted in Figure 1, upfixes use a diverse range of images and symbols to convey their meaning, often with varying types of reference (Forceville, El Refaie, & Meesters, 2014). Some upfixes involve symbols with fixed meanings, such as hearts or exclamation marks, which retain their meaning even away from a face. Other upfixes derive from idiomatic verbal expressions, such as "seeing stars" resulting in stars twirling above characters' heads to show dizziness (Cohn, 2016). Still others derive meaning through metaphors (Forceville, 2011; Lakoff & Johnson, 1980), often altering iconic representations. For example, gears turning above the head indicate thinking, which invokes the metaphors that the MIND IS A MACHINE and MOVEMENT IS PROGRESS (Cohn, 2010), while storm clouds meaning a bad mood rely on a metaphor of WEATHER AS AN EMOTIONAL FORCE (Shinohara & Matsunaka, 2009). Thus, specific upfixes use various methods to derive meaning; though, as visual signs related to emotional/cognitive states, upfixes in general may involve metaphorical understanding due to their proximity to the head and face.

Some prior work has argued that comic readers do interpret upfixes as conveying emotional meanings, beyond the facial expressions they accompany (Feng & O'Halloran, 2012). We might think of this as an item-based "lexical" theory, whereby the upfix results in the retrieval of a specific stored meaning. In one study, Ojha (2013) asked participants to interpret four different types of upfixes (spirals, spikey lines, twirls, and sweat drops) placed above faces showing neutral expressions. When choosing amongst a list of possible interpretations for their expected meanings (anger, surprise, confusion, and agitation), participants identified a variety of emotions, but most frequently chose the same two meanings (surprise: mean = ∼38%, confusion: mean = ∼38%) regardless of the specific upfix. While these results support that upfixes significantly contributed to the interpretation of emotion, they contrasted the idea that certain upfixes carried specific meanings. In a second study, participants were given this same list of particular emotions and were forced to choose an upfix paired with a neutral face which best represented that emotion. Here, interpretations more consistent with the specific upfixes appeared (max = 53%), though several responses were still found for each upfix. In addition, there was surprisingly no interaction between participants' interpretations and their comic reading expertise.

In theoretical work, we have argued that upfixes go beyond this item-based lexical account (Cohn, 2013). Rather, these representations belong to a broader class that uses an abstract *lexical schema* stored in the long-term memory of individuals who have acquired this visual vocabulary (prototypically, comic readers). This affixation schema specifies that some type of graphic representation (an upfix) is placed in an upward relation to a head, and this

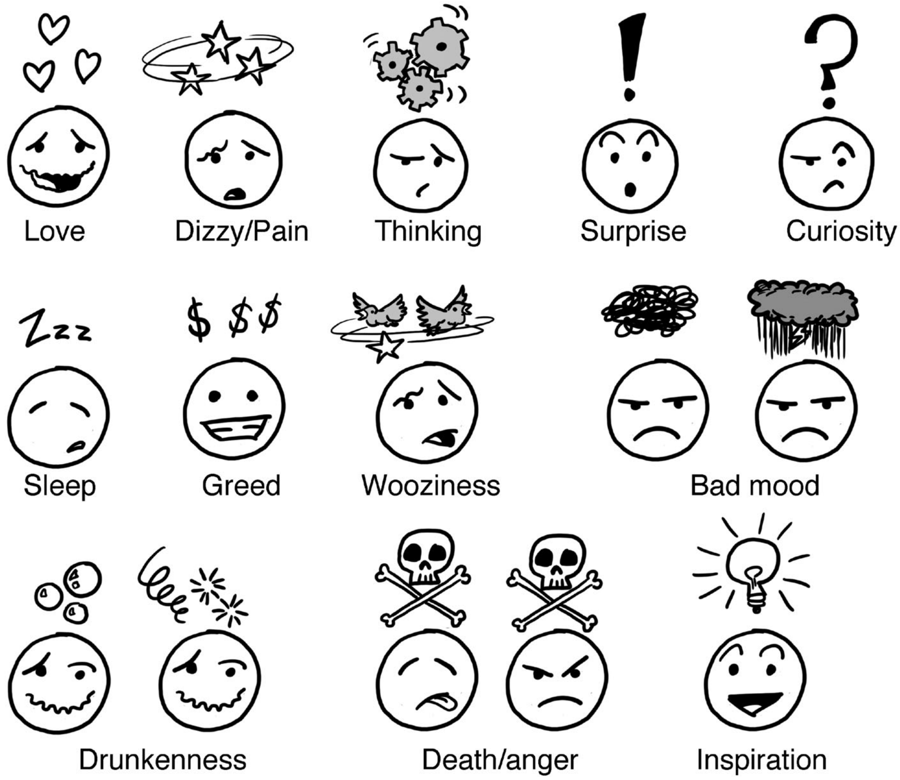

Figure 1. Conventionalised "upfixes" from the visual vocabulary used in comics.

juxtaposition results in a holistic meaning integrating those parts, usually related to emotional or mental states. While conventionalised upfixes are stored in memory (similar to the item-based theory), this abstract schema is "semi-productive", allowing for the creation of novel upfixes that might use this broader pattern.

This account is structurally analogous to construction grammar-based models of morphological schemas in verbal language (Booij, 2010; Jackendoff, 2002). Here, affixes have schematic structure that allows variable and productive usage, such as the English suffix—*ness*, which has a structure of $[[X]_A\text{-ness}]_N$ which links to a semantics of "the property/state of A" (Booij, 2010). As a schema stored in memory, this allows for both memorised instances of this affix in conventionalised words (*awareness*, *happiness*), but also productive, novel forms (*comic-ness*).

Upfixes would thus be similar, operating with variability for both graphic sign and facial expressions, but always in a face–upfix juxtaposition. Following Booij (2010), we may articulate an upfix schema as:

$$[[\text{Emotion Expression}_i]_{\text{Head}} \text{ below } [x]_{\text{Upfix}}]_j \leftrightarrow [\text{Type of Emotion}_i]_j.$$

The left side of this schema specifies its form (the head is *below* the upfix), while the right side specifies the construed holistic meaning. Within the left side, the brackets marked *Head* and *Upfix* are slots, which can be filled with appropriate representations. The head takes on a face with an "emotional expression" which is presupposed to exist as an independent visual sign; indeed, emotional faces do appear without upfixes. In contrast, upfixes cannot exist independently, but rather are morphemes bound to their root (the head).

This schema also specifies constraints between upfix and head. First, upfixes are restricted to some space above the head, as specified by the relation "below" in the left side of the schema. Thus, a lightbulb above the head to indicate inspiration would make less sense if placed beside the head (first row of Figure 2(b)). Second, upfixes are also constrained by a particular "agreement" between the facial expression and the graphic sign. This agreement is specified in the schema by mapping the emotions specified by the head/face (i) in combination with an upfix (j) to the holistic emotion on the right side of the schema. For example, the lightbulb must accompany a happy face, and the meaning should be stranger if placed above a

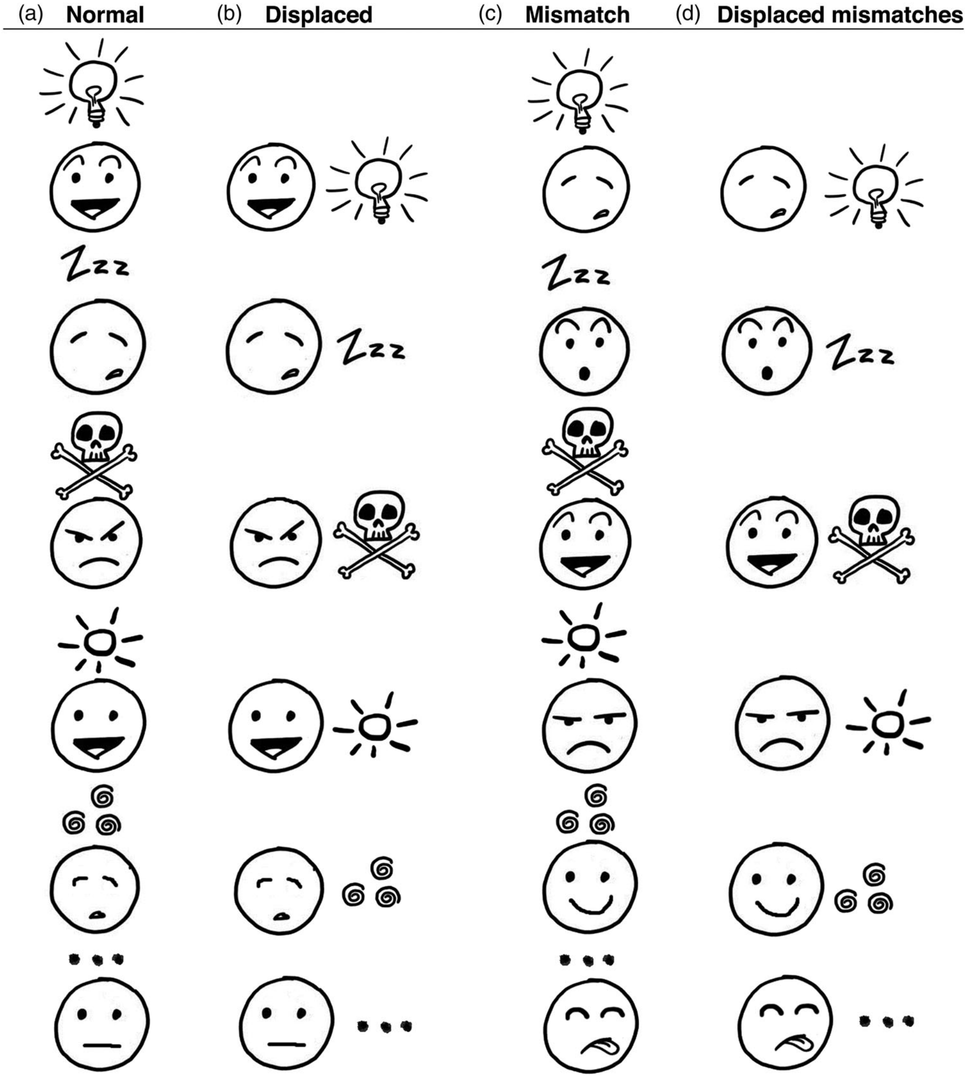

Figure 2. Examples of constraints on upfixes (lightbulb, Zzz, skull-and-crossbones, sun, spirals, ellipses): (a) normal upfixes positioned above the head and agrees with its facial expression. (b) Upfixes displaced from their canonical position above the head. (c) Upfixes with faces that mismatch their meaning. (d) Displaced mismatching upfixes.

sleeping face (Figure 2(c)). Finally, we might also hypothesise a relationship between these constraints, where an even more strained interpretation would arise from violating both constraints, such as when a lightbulb appears beside a head and with a sleeping face (Figure 2(d)).

Agreement between face and upfix should be motivated by item-specific constraints carried by the upfix. For conventional upfixes, this relationship would be stored in memory as a particular item-specific instantiation of the schema above, similar to relations stored in specific lexical items of language (e.g. Booij, 2010). For example, a lightbulb as an upfix may carry with it specifications for accompanying a happy or inspired face, while storm clouds would license being associated with a sad face (and, potentially, vice versa). That is, meaning does not come from the face or upfix alone, but out of their combination. We posit that comprehenders abstract across the observations of these item-specific instances to form the abstract upfix schema. For novel upfixes, this relationship may be construed from the semantic associations and/or metaphoric implications of the graphic signs (Forceville, 2011; Lakoff & Johnson, 1980), but are fairly constrained by the facial expression of the head. For example, an upfix of a sun is fairly unconventional, yet may involve a metaphor of WEATHER AS AN EMOTIONAL FORCE (Shinohara & Matsunaka, 2009) along with LIGHT AS GOOD (Forceville &

Renckens, 2013). Thus, for a sun, we would expect that a happy face would be more congruent (Figure 2(a), fourth row) than an upset face (Figure 2(c)). However, this novel, construed relationship should be less strong than meanings already entrenched in memory. Mismatches between faces and conventional upfixes, which are stored in memory, should thus have a larger impact than between unconventional pairs.

This principled relationship between face and upfix may explain the variety of interpretations found in Ojha's (2013) studies. Since those experiments used neutral faces, the upfixes had no specific "bound" relationship to their accompanying facial expressions. Essentially, these upfixes "disagreed" with their faces, although this may have been a "weak" disagreement because the faces used neutral expressions rather than offered conflicting emotions. If this were the case, we would expect that more consistent interpretations would arise from upfixes that agree with their facial expression than those that mismatch their facial expression.

Another theory has argued against upfixes as using an abstract schema constrained by placement and agreement with a face, instead favouring the idea that all upfix meanings are computed by semantic construal on a contextualised basis. For example, Bateman and Wildfeuer (2014) claim that a lightbulb's meaning merely "elaborates" on the implied mental state of the facial expression using a "defeasible discourse interpretation", and that its proximity as "near" a head is enough. Because this construal views all face–upfix relationships as possible, mismatches would be just as interpretable as matches, given context. In the case of an upset face with a lightbulb upfix, they claim a potential interpretation would be that the person is both unhappy *and* inspired. Note, however, that this example does not integrate the signs into a holistic meaning (as in the proposed schema above), but rather maintains both independently (and, notably, offers no explanation for how lightbulbs mean inspiration, implying an item-based stored meaning). Such a purely dynamic semantic approach implies that no privileged, entrenched relationship exists between specific upfixes and their facial expressions, and that mismatches would be equally acceptable as matches, with both undergoing the same process of construal. Also, because such construal should operate dynamically on all face–upfix pairs, it would predict no difference between upfixes that are conventionalised from those that are novel.

Given these precedents, we therefore sought evidence that comic readers store this abstract "upfix schema" in their long-term memory—beyond just storing item-based meanings—and that restrictions on location and agreement constrain the comprehensibility and interpretation of these combinatorial signs. In two experiments, participants were presented with images of faces that accompanied conventional and unconventional upfixes. Our normal images used upfixes that both agreed with the face and were placed above the head. These upfixes were then manipulated to either mismatch the face (Experiment 1) and/or be moved to the side of the face (Experiment 2), as in Figure 2. In Experiment 1, participants rated these images for their conventionality and the degree to which the face and upfix "went together". In Experiment 2, participants rated these images for how "easy they were to understand". In both experiments, participants also offered freely given interpretations for upfix meanings.

2. Experiment 1: conventionality and agreement

Our first experiment sought to confirm that certain upfixes are more conventionalised than others, and that upfixes have a relationship to the face they accompany. Ultimately, "conventionality" may fall on a continuum based on relative proportions within a broader visual language. Such an account would be consistent with both an item-based theory and the constructional schema theory (Booij, 2010; Jackendoff, 2002), and these proportions could be explored via widespread corpus analyses (Forceville, 2011; Newton, 1985). In the absence of this corpus data, we here use a binary split to elicit distinctions from participants' own ratings. We therefore presented participants with upfixes that we expected to be more or less conventional, which either matched or mismatched their facial expressions. Participants rated them for how familiar they were and for how well they "went together" with their accompanying face, and then provided an interpretation for their meaning.

If the item-based theory is correct that these forms are merely stored as individual tokens in memory, conventional upfixes should be rated as easier to understand than unconventional ones and matching face–upfix relations should be easier to understand than mismatches. However, mismatches should not vary based on conventionality, since they would be equally unfamiliar. A stronger item-based view might also posit that mismatches would not be worse than matches, since the upfix alone carries a specific meaning regardless of facial expression (e.g. Ojha, 2013).

In contrast, a dynamic process of construal operating face–upfix relations (Bateman & Wildfeuer, 2014) predicts no modulation based on conventionality or face–upfix (mis)matches, since all relations should be feasible through a dynamic process of interpretation. While our stimulus presentation did not embed these elements in a narrative context, isolated forms should provide the most open interpretative possibilities, since participants were free to invent their own feasible contexts.

Finally, the schema theory predicts that conventional upfixes will be rated as higher than unconventional ones, and that matching face–upfix relations would be better than mismatches. This difference between matching and mismatching face–upfix pairs should occur in both conventional *and* unconventional instances, since the schema operates over both types. However, this difference should be larger for more conventional face–upfix pairs, which should gain an advantage from being entrenched in memory.

2.1. Methods

2.1.1. Stimuli

We created 32 face–upfix pairs, comprised of what we categorised as 16 conventional upfixes and 16 unconventional upfixes. Our unconventional upfixes were images used in a novel way that could have a logical semantic association (rainbows, marijuana leaf), a fixed meaning (peace sign), or more abstract shapes that should contain no overt meaningful associations (plus signs, circles, triangles). Conventional upfixes were: hearts, stars, gears turning, an exclamation mark (!), a question mark (?), Zzzzs, dollar signs ($), birds and stars, storm clouds, bubbles, skull-and-crossbones, light bulb, twirl and stars, scribble, halo, and music notes (a selection of which appear in Figures 1 and 2). Unconventional upfixes were: triangles, a flame, a marijuana leaf, a rainbow, a four leaf clover, clouds, a single large water droplet, a fork and knife, Xs, plus signs (+), spirals, a peace sign, a sun, ellipses (…), sparkles, and circles.

Matching upfixes had an emotion that agreed with the meaning of the face (Figure 2(a)), while mismatching upfixes altered the emotional expression of the face so that it disagreed with the upfix (Figure 2(c)). Face–upfix pairs were distributed into four lists counterbalanced such that each participant viewed each upfix only once, and each list presented stimuli in a randomised order. These lists also rotated through 25 filler face–upfix pairs of varying degrees of conventionality.

Because meaning might vary based on the relationship between upfix and face, we used a variety of different reference types throughout stimuli posited as conventional and unconventional. “No meaning” signs had no intrinsic meaning when separated from the upfix, such as triangles or scribble. “Fixed” meanings had a symbolic meaning outside of their use as upfixes, such as hearts (regardless of whether their origins may have been metaphoric or metonymic). “Metaphoric” meanings used underlying mappings between domains, such as lightbulbs or gears (Forceville, 2011; Lakoff & Johnson, 1980), and finally “associative” meanings may have had intrinsic and/or metonymic meaning, which changed when acting as an upfix, such as spiraling birds. Repetition of emotional expression (such as happy or angry/grumpy) used several different characters’ faces, so as not to repeat the same face multiple times.

2.1.2. Participants

Eighty-seven participants (55 males, 32 females, mean age: 34.3) completed our experiment via an online survey. All participants gave their informed written consent via a digital signature. Prior to experimentation, all participants filled out the “Visual Language Fluency Index” (VLFI) questionnaire used to assess their expertise at the visual language of comics by asking about the frequency with which they read various types of visual narratives (comic books, comic strips, graphic novels, Japanese comics, etc.) and drew comics, both currently and while growing up. These ratings were then incorporated into a formula that calculated a “VLFI score” that has been shown to correlate significantly with both behavioural and neurocognitive measures (Cohn et al., 2012; Cohn & Wittenberg, 2015), including manipulations to “visual morphology” (Cohn & Maher, 2015). An idealised average along this metric would be a score of 12, with low being below 7 and high above 20. Participants’ mean fluency was high, 22.1 ($SD = 9.7$), though overall they had a wide range of expertise (range = 2.13–45).

2.1.3. Procedure

Participants completed the experiment using a web browser. On each page of the survey, participants were presented with a single face–upfix pairing and were asked to rate how familiar it was (1 = not familiar, 7 = very familiar), to rate how well the face and upfix "went together" (1 = do not belong together, 7 = belong together), and to provide an interpretation of the overall meaning.

2.1.4. Data analysis

For both ratings of familiarity and matching, we averaged across participants' ratings for each type of upfix and then calculated the mean rating for each condition for each participant, collapsing across items. Our analysis of participants' freely given interpretations of the upfixes categorised all participant responses into one of 27 different semantic/emotional categories. These categories were assigned based on knowledge of conventionalised upfixes and graphic depictions of emotional facial expressions. They included: Happy, angry/ grumpy, peaceful, love, dizzy/dazed, pain, surprise, curious/ unsure, sleepy/tired, drunk, death, greed, thinking, lucky, high, anxious/nervous, daydreaming, angelic, inspired, hungry sad/depressed, singing, confused, afraid, sick/ill or pensive/speechless. An additional category of "iconic explanation" was given to interpretations that stated upfixes were an iconic object, such as storm clouds being actual clouds and rain. A final label of "other" was used where interpretations were ambiguous or unclear. Coders carried out categorisation of participant responses while blind to the pictorial stimuli. For each stimulus, we calculated the most frequent (mode) response across all participants, and then calculated whether each participant agreed (1) or disagreed (0) with that mode response. For each participant, we then calculated their mean agreement for mode interpretations for each condition, collapsing across items.

Means for ratings and interpretations were analysed using 2 (Conventionality: conventional vs. unconventional) × 2 (Matching: match vs. mismatch) repeated-measures ANOVAs. In the case of statistical interactions, follow-up *t*-tests were used to analyse pairwise relations, with a Bonferroni correction for multiple comparisons. Finally, to investigate the role of comic reading frequency on participants' assessment of the stimuli, differences between mean ratings and between interpretations were then correlated with each individual's VLFI score using Pearson's correlation set to an alpha level of .05.

2.2. Results

2.2.1. Conventionality ratings

Participants' "familiarity" with face–upfix pairs distinguished conventional and unconventional representations, and matching from mismatching ones. We found significant main effects for Conventionality and Matching, and a significant interaction between them (Table 1). As depicted in Figure 3, conventional face–upfix pairs were rated as more familiar than unconventional ones, and matches were more familiar than mismatches. The effect of matching was larger in conventional than unconventional pairs, leading to the statistical interaction, but all levels were significantly different (all ts > 2.9, all ps < .005).

A positive correlation appeared between VLFI scores and the difference between familiarity ratings for conventional and unconventional mismatches, $r(85) = .243$, $p < .05$. This suggests that greater comic reading experience leads to larger differences between conventional and unconventional mismatching face–upfix pairs. Conversely, a negative correlation between VLFI scores and the difference between conventional matches and mismatches, $r(85) = -.215$, $p < .05$, implies that participants with greater comic reading experience had less disparity between their familiarity ratings of these types.

2.2.2. Matching ratings

We next analysed the ratings for how well the face and upfix "went together". Again, significant main effects appeared for Conventionality and Matching, and their interaction (Table 1). As in Figure 3, this arose because upfixes that matched their faces were rated higher than those that mismatched, while conventional matches were rated higher than unconventional ones. However, although conventionality had a large effect in matching pairs $t(86) = 14.6$, $p < .001$, mismatching face–upfix pairs did not differ in this respect ($p = .120$).

2.2.3. Interpretations

Our final analysis examined the rate at which participants volunteered the most frequent (modal) interpretation of upfixes, where again we found main effects of Conventionality and Matching, and

Table 1. Results of ANOVAs comparing ratings and rates of most frequent (mode) interpretations face–upfix pairs in Experiment 1.

	Familiarity ratings			Match ratings			Mode interpretations		
	F-value	MSE	η_p^2	*F*-value	MSE	η_p^2	*F*-value	MSE	η_p^2
Conventionality (C)	212.2***	1.23	0.712	49.6***	1.18	0.366	32.7***	0.052	0.276
Matching (M)	77.2***	1.05	0.473	390.4***	0.89	0.819	138.1***	0.059	0.616
C * M	22.68***	0.88	0.231	159.2***	0.62	0.649	9.24**	0.051	0.097

df = 1.86.
****p* < .001.
***p* < .01.
**p* < .05.
^*p* < .1.

an interaction between them (Table 1). As in Figure 4, this statistical interaction arose because upfixes that matched their face were interpreted with more agreement than those that mismatched, while conventional matches were agreed on more than unconventional ones. However, only a trending significance suggested that conventionality differed between interpretation of mismatches, $t(86) = 1.75$, $p = .083$.

2.3. Discussion

This study sought to establish that certain upfixes would be more familiar (i.e. conventional) to participants than others, and that upfixes have a preferred relationship to the facial expressions that they accompany. First, participants clearly distinguished between their recognition of conventional and unconventional upfixes in both types of ratings. Participants rated conventional upfixes as more familiar than unconventional upfixes, while matching faces were rated as more familiar than mismatches, regardless of conventionality. Altogether these ratings suggested that certain upfixes are more conventional than others, which includes particular agreement relationships between upfixes and faces. These results of familiarity are consistent with prior research showing that the understanding of visual morphology, including upfixes, aligns with their frequency in comics (Newton, 1985).

These findings were repeated in the matching ratings. Conventional matching upfixes were rated as "belonging with" their face more than any other type. However, mismatches were also rated worse than matches even within unconventional pairs, suggesting that even novel upfixes adhere to some degree of agreement with their faces. This is a critical test of the processes underlying

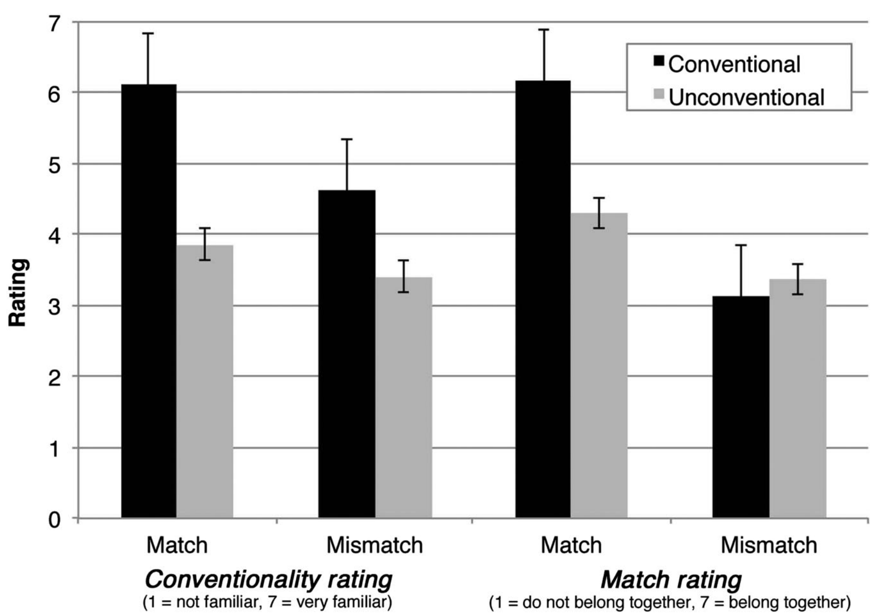

Figure 3. Conventionality and matching ratings given to conventional and unconventional face–upfix pairs that did or did not match in facial expression. Error bars depict standard error.

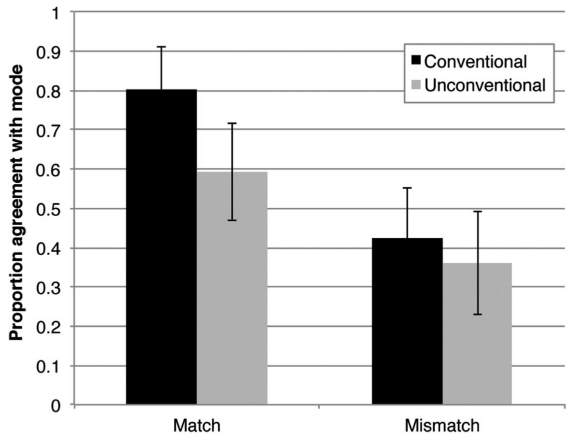

Figure 4. Proportion of participants that agreed with the most frequent (mode) interpretation of the meaning of face–upfix pairs.

interpretation of visual morphemes. Because the agreement between face and upfix operates on both conventional and unconventional upfixes, it suggests that they are understood via an abstract pattern, not memorisation of specific items. The familiarity ratings further support this interpretation, showing the same difference between matching and mismatching face–upfix pairs in both conventional and unconventional cases. Nevertheless, in both types of ratings the difference between matches and mismatches was greater for conventional than unconventional face–upfix pairs, consistent with the idea that storing conventional face–upfix pairs in memory would give them an advantage over unconventional representations.

Participants' freely given interpretations further reinforce these findings, which found high consistency for the meanings of both conventional and unconventional matching upfixes. As expected though, meanings for conventional upfixes were agreed upon more than unconventional ones. In contrast, the consistency of interpretations to mismatches was far lower, and only trended towards differing based on conventionality. This supports that mismatching upfixes were harder to interpret than those that agreed with their faces, which would support the idea that specific upfixes carry particular expectations about the way they should relate to faces.

It is also worth noting that participants overtly noted the mismatch between face and upfix in their responses. They stated that mismatches were "incongruous", "not vernacular" or that "the face doesn't match", "the expression looks wrong for this", or "I can't repair the semantics here". Participants also made corrections, such as for the mismatching face to a Zzz upfix, stating "the eyes should be closed" (i.e. sleeping), or that skull-and-crossbones "should not be paired with a happy face" and others. In mismatching contexts, participants often used certain strategies that avoided providing a holistic interpretation. They instead would either default to interpreting them on the basis of the conventionalised meaning for either upfix or face separately, would list both the conventionalised meanings of both face and upfix separately with no attempt at combining them, or would provide causative explanations treating upfixes as iconic elements like "they are happy because it's raining".

Because participants' strategies often provided meanings for upfixes and faces independently in mismatching contexts, it supports that both parts of this combination carry distinct meanings. However, because these responses also convey an unwillingness to combine them, it contrasts with the idea that participants will dynamically construe meanings for mismatching face–upfix pairs, without recognising them as incongruous (Bateman & Wildfeuer, 2014). Rather, participants both implicitly and explicitly stated their incongruity.

Similar responses arose for unconventional upfixes. Participants pointed out the non-entrenched character of these upfixes with comments like "the meaning is clear enough, but it's not effective", "it doesn't really work", or that it was "unconventional … [with] no specific meaning". Along with the different ratings and rates of interpretations for conventional and unconventional upfixes, these results point toward a constrained abstract pattern stored in memory, not a generalised process of construal or storage of specific items alone.

3. Experiment 2: conventionality, agreement, and placement

In Experiment 1, we established that participants recognise differences in conventionality between upfixes, and that they privilege certain relationships between upfixes and faces. In our second experiment, we therefore investigated the role of upfix placement in combination with its agreement to the face. As stated above, we hypothesised that displacing an upfix next to a face will result in lower comprehensibility than when above a head. Such a result should not occur if upfixes merely need to have a proximity of being "near" a face in order for construal of their relationship (Bateman & Wildfeuer, 2014), with no specific constraint on being above the head. Moreover, we predicted that this constraint of placement should interact with the constraint on agreement, such that a dual violation should be worse than violating a single dimension. Thus, in contrast with Experiment 1, where we investigated aspects of conventionality or internal structure (agreement), in Experiment 2 we examined how well participants understood the meanings of these varied upfixes. We therefore asked participants to rate upfixes for their overall comprehensibility and again interpret their meanings.

3.1. Methods

3.1.1. Stimuli

We again used the stimuli from Experiment 1 with conventional and unconventional upfixes that either matched or mismatched their accompanying facial expressions. However, we added an additional factor of Placement, whereby the upfix was either above the head or displaced to the side of the head. This created eight possible types: normal upfixes had an emotion that agreed with the meaning of the face and were located above the head (Figure 2(a)). Displaced upfixes were moved from above the head to beside the head (Figure 2(b)). Mismatching upfixes altered the emotional expression of the face so that it disagreed with the upfix (Figure 2(c)). Finally, displaced mismatches both moved the upfix to the right side of the head and altered the emotion so that the face disagreed with the upfix (Figure 2(d)). These manipulations were applied to both conventional and unconventional face–upfix pairs.

Altogether, these manipulations yielded a 2 (Conventionality: conventional vs. unconventional) × 2 (Placement: above head vs. beside head) ×2 (Matching: match vs. mismatch) experimental design. Stimuli were counterbalanced using a Latin Square design into four separate lists each containing 32 face–upfix pairs, such that each participant viewed each type of upfix only once. We then created packets containing these stimuli, which presented them in a randomised order.

3.1.2. Participants

Our experiment was taken by 70 participants (37 males, 33 females, mean age: 21.5) from the UC San Diego community who gave their informed written consent according to the guidelines of the UCSD Human Research Protections Programme. Participants had an average fluency with comics, with a mean VLFI score of 15.7 (SD = 9.1, range = 1.75–41.25).

3.1.3. Procedure

Participants were given packets that contained the various face–upfix pairings. Beneath each graphic was a row of numbers from 1 to 7 where participants circled the rating for how easy the meaning was to understand (1 = very difficult, 7 = very easy). Below this rating, participants were given a line where they were asked to write their interpretation of the images. The experiment took participants roughly 5 min to complete.

3.1.4. Data analysis

We used the same methods of data analysis for ratings and interpretations as in Experiment 1. Ratings and interpretations were analysed using 2 (Conventionality) × 2 (Placement) × 2 (Matching) repeated-measures ANOVAs, followed by *t*-tests to analyse pairwise interactions between conditions. Differences between mean ratings and between per cent of agreement for interpretations were then correlated with each individual's VLFI score using Pearson's correlation set to an alpha level of .05.

3.2. Results

3.2.1. Ratings

Analysis of participants' ratings found main effects of Conventionality, Placement, and Matching (see Table 2). A significant interaction appeared between Conventionality and Matching, and a trending interaction between Conventionality, Placement, and Matching ($p = .089$). No two-way interactions were found between Conventionality and Placement ($p = .958$), or Matching and Placement ($p = .381$).

To break down the three-way interaction, Bonferroni-corrected *t*-tests were used to examine pairwise relations between conditions. There is a main effect of conventionality because conventional face–upfix pairs were generally more understood than unconventional ones (Figure 5(a)). However, this is qualified by interactions between Conventionality and Matching, and a trending three-way interaction between Conventionality, Placement, and Matching. Looking at the pairwise relations, conventional face–upfix pairs were rated higher than unconventional

Table 2. Results of ANOVAs comparing ratings and rates of most frequent (mode) interpretations face–upfix pairs in Experiment 2.

	Comprehensibility ratings			Mode interpretations		
	F-value	MSE	η_p^2	*F*-value	MSE	η_p^2
Conventionality (C)	33.14***	1.03	0.324	10.9**	0.03	0.135
Matching (M)	180.6***	1	0.724	52.3***	0.06	0.427
Placement (P)	41.49***	0.87	0.376	0.334	0.04	0.005
C * M	23.93***	1.15	0.258	15.3***	0.07	0.179
C * P	0.003	0.56	0	3.27^	0.03	0.044
M * P	0.776	0.89	0.011	3.16^	0.06	0.043
C * M * P	2.97^	1.16	0.041	4.96*	0.05	0.066

df= 1.69.
****p* < .001.
***p* < .01.
**p* < .05.
^*p* < .1.

ones when matching (normal, displaced), all *t*s > 3.9, all *p*s < .001, but not when mismatching (mismatch, dual), all *p* > .247. Matching face–upfixes pairs (normal, displaced) were rated higher than mismatching pairs (mismatches, displaced mismatches), all *t*s > 4.1, all *p*s > .001, regardless of conventionality. The main effect of Placement arose because normal face–upfix pairs were rated higher than displaced ones (all *t*s > 2.3, all *p*s < .05), regardless of conventionality. Mismatches were also rated higher than displaced mismatches, a trending difference for conventional face–upfix pairs, $t(69) = 1.8$, $p = .08$, but statistically significant for unconventional face–upfix pairs, $t(69) = 4.02$, $p < .001$.

We found no significant correlations between comprehension ratings and VLFI scores.

3.2.2. Interpretations

Our second analysis focused on participants' interpretations of the face–upfix pairs by assessing the rate at which participants agreed with the most frequent (mode) responses for their meanings. As in Table 2, we found main effects for Conventionality and Matching, but not Placement ($p = .728$). Significant interactions appeared between Conventionality and Matching and Conventionality, Matching, and Placement, while trending interactions appeared between Conventionality and Placement as well as Matching and Placement.

As depicted in Figure 5(b), the largest disparity in consistency of interpretations came when there was disagreement between face and upfix, though primarily for conventional upfixes. Our follow-up pairwise analyses showed that interpretations for conventional mismatching face–upfix pairs (mismatch, displaced mismatch) were agreed upon less than for matching pairs (normal, displaced), all *t*s > 5.6, all *p*s < .001. Rates of agreement for unconventional normal upfixes did not differ from unconventional mismatches or displaced mismatches, all *p* > .441, although unconventional displaced upfixes had more agreed upon interpretations than unconventional mismatches or displaced mismatches (all *t*s > 1.9, all *p*s < .001). Conventionality most strongly influenced the interpretations of normal face–upfix pairs, with conventional pairs rated higher than unconventional ones, $t(70) = 8.9$, $p < .001$. However, trending differences also appeared for displaced, $t(70) = 1.96$, $p = .055$, and mismatching face–upfix pairs, $t(70) = -1.85$, $p = .068$, but not displaced mismatches, $p = .548$. Placement had little influence on interpretations, with no difference appearing between upfixes moved to the side of the head (displaced, displaced mismatch) from those that were above the head (normal, mismatch), regardless of conventionality. The only exception to this was the higher agreement on interpretations for unconventional displaced and normal upfixes, $t(70) = -2.6$, $p < .05$.

To further explore the influence of participants' comic reading experience on their interpretations of face–upfix pairs, we compared the differences between mode interpretations and VLFI scores. Positive correlations were found between VLFI scores and the difference between interpretations of conventional and unconventional displaced face–upfix pairs, $r(69) = .321$, $p < .01$, and conventional displaced face–upfix pairs and displaced mismatches, $r(69) = .275$, $p < .05$, suggesting a larger difference in agreement of interpretations by participants with greater comic expertise. A negative correlation between VLFI scores and the difference between rates of agreement for the interpretations of unconventional displaced and mismatching face–upfix pairs, $r(69) = -.276$, $p < .05$, suggested that

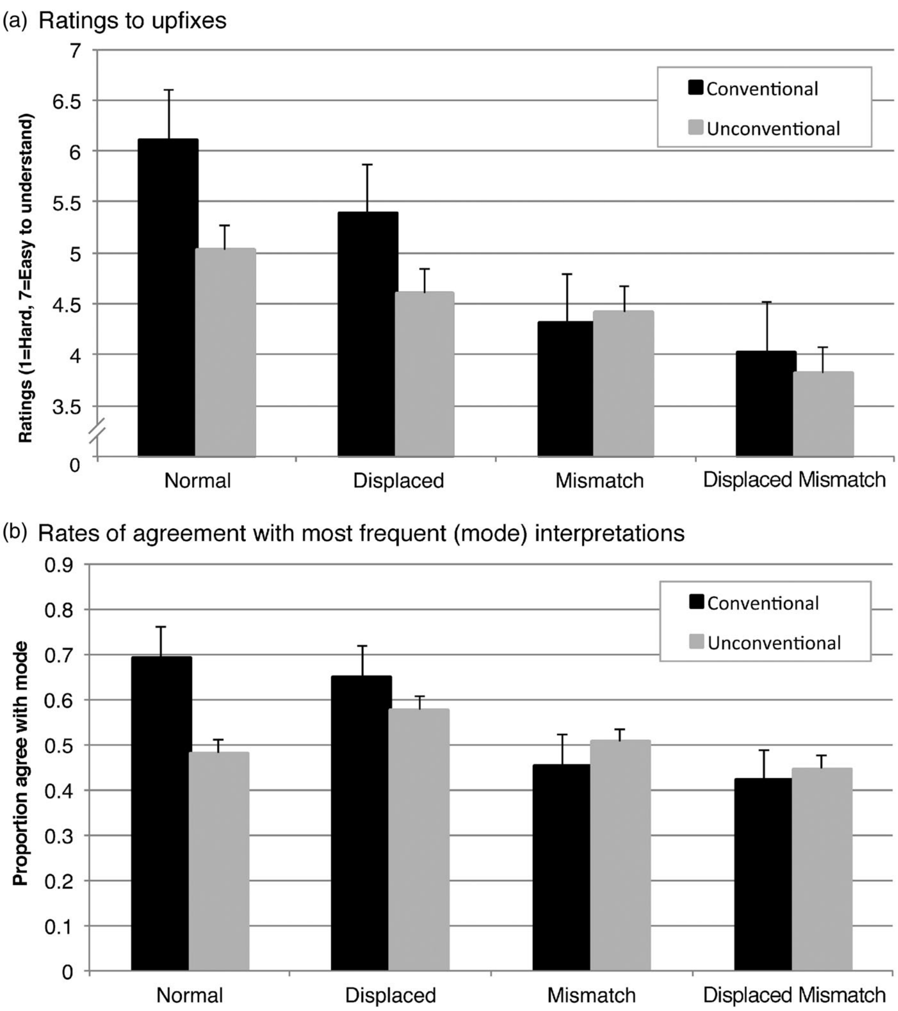

Figure 5. Participants' (a) comprehensibility ratings and (b) rates of agreeing with the most frequent (modal) interpretations for images where upfixes were either in their normal placement and matched the face, were displaced beside the head and/or mismatched the emotion of the face. These results grouped stimuli as conventional—appearing commonly in comics—or unconventional—upfixes created explicitly for this experiment. Error bars depict standard error.

participants agreed less on their meaning when they had greater comic reading expertise.

3.3. Discussion

This experiment sought to confirm that, not only are upfixes constrained by their relationship to a face, but also to their placement relative to the face. We therefore placed both the conventional and unconventional upfixes from Experiment 1 above and next to their accompanying faces, along with manipulations so that they either matched or mismatched the facial expressions.

First, we reconfirmed that upfixes have a particular relationship with their faces. Participants rated upfixes that matched their faces (normal, displaced) as more comprehensible than those that mismatched (mismatch, displaced mismatch). In addition, placement of the upfix also mattered. Upfixes placed above the head (normal, mismatch) were rated higher than those displaced next to the head (displaced, displaced mismatch), again for both conventional and unconventional upfixes. Furthermore, violating both an upfix's placement and agreement with a face created a combined impact: displaced mismatches were worse than both mismatches and displaced face–upfix pairs of both types of conventionality. Because the placement of an upfix influences participants ratings of comprehensibility, it provides evidence against a view that proximity alone is sufficient without consequences on congruity (Bateman & Wildfeuer, 2014). Furthermore, because these effects occurred for both conventional and unconventional upfixes, it implies an abstract pattern underlying their comprehension, not simply item-based memorisation.

Nevertheless, the matching relationship between face and upfix appeared to be more impactful on interpretations than placement. Participants' agreed more on interpretations for upfixes that matched their faces than those that mismatched, regardless of their location. This suggests that both the face and the upfix contribute meaningful information, which licenses the types of elements they can accompany. This information allows faces and upfixes to form a combinatorial meaning through a prototypical relative positioning above the head. However, more meaning may be recoverable when displacing the upfix than if those component parts do not agree, since the underlying semantics are still recoverable, despite their awkward proximity. Agreement between visual morphemes is thus a stronger constraint on the semantics of upfixes than placement, though both factor into the overall meaning (as indicated by the differences in ratings across all manipulations, especially displaced mismatches).

Despite this relative lesser influence by upfix location, the positional information tested here was fairly restricted: for placements above versus beside the head, position appears to contribute little dissociable interpretation so long as the elements match each other. However, other locations may carry more semantic weight. As discussed, twirls above the head mean something different than behind a body (Forceville, 2011), and stars in the eyes differ in meaning from those above the head (Cohn, 2013). Comparison between visual morphemes where the positions carry meaningful contrast may therefore yield different results.

4. General discussion

These experiments sought evidence for an abstract combinatorial schema within visual representations. We examined whether "upfixes"—the graphic signs that often float above character's heads—are constrained by their placement above the head and by their agreement with a face's emotion. Overall, we found that both constraints impacted their ratings and interpretations—no matter the conventionality—providing support for the idea that these graphic signs use an abstract schema.

Across both experiments mismatches between upfixes and their preferred facial expressions resulted in lower ratings of familiarity, "belonging together" and comprehensibility. Such mismatches also resulted in less consistent interpretations of upfix meanings. These results suggest that a preferred "agreement" between upfix and faces constrains their meaning. Experiment 2 showed further that altering the location of upfixes also impacts their comprehensibility, but not as much as agreement. Displaced upfixes were rated as more comprehensible than those that mismatched their accompanying face. Yet, ratings to displaced mismatches were lower than mismatches alone, suggesting that manipulation of both factors creates a compounded effect of violating both constraints.

Crucially, in both experiments, manipulations of agreement and placement affected the ratings of both conventional *and* unconventional upfixes. If upfixes were simply stored as item-based instances and/or relied only on contextually defined construal, then we would expect unconventional upfixes to be rated the same no matter their agreement or placement relative to a face, since these novel instances would have no instantiation in memory. Rather, ratings to unconventional upfixes—be it for familiarity, matching, or comprehensibility—displayed the same pattern of decrement as to conventional upfixes. This suggests that constraints on matching and placement applied to the comprehensibility of upfixes beyond individual conventionalised signs, but rather reflects an abstract schema stored in memory.

Nevertheless, conventionality did factor into the understandability of upfixes. In both experiments, conventional upfixes were rated as more familiar, as more meant to "go together" and as easier to understand than unconventional upfixes, but only for those that matched their faces. Conventional and unconventional upfixes that mismatched their face were rated as equally understandable. Furthermore, the difference between ratings of matches and mismatches were greater for conventional than unconventional face–upfix pairs. Participants' interpretations of these upfixes reinforced these results. The placement of the upfix mattered less in participants' interpretations of upfix meanings than agreement. In both experiments, upfixes that matched their faces were more consistently interpreted than mismatches.

For mismatches, conventionality had little influence on interpretations. In both experiments, the rates of interpretation for mismatching face–upfix pairs hovered around 45%, which, it is worth noting, is not a marginal number for a freely generated response. This consistency may indeed reflect a striving for dynamic construal (Bateman & Wildfeuer, 2014) that seeks to garner meaning out of mismatching parts. If so, such rates of interpretation are significantly lower than those where the elements have a recognisable relationship, whether fully conventionalised or novel. Thus, such a construal process may be engaged in lieu of a privileged, conventional relationship. Despite this, it is worth pointing out that participants' strategies for mismatches generally used non-holistic interpretations—that is, not attempting to integrate conflicting faces and upfixes.

Because conventionality had an influence overall though, these results suggest that upfixes do have item-specific constraints, despite tapping into an abstract schema. That is, upfixes on their own do not determine the meaning by mere placement above a head, but rather each upfix carries specifications for how it should contextually relate to an accompanying face. These specifications are more advantageous for conventional upfixes, which are stored in memory, than unconventional ones, which are less entrenched. These results may therefore inform why interpretations of upfixes may have been less forthcoming above neutral faces, as in Ojha's (2013) study: neutral faces would "mismatch" their upfixes, and thus yield more variability of interpretations than with matching faces.

It is also worth noting that, methodologically, the present study elicited *freely given* responses from participants. Yet, these interpretations of normal conventional upfixes (Experiment 1: 80%, Experiment 2: 70%) were far greater than the highest rates of interpretation for all upfixes in Ojha's (2013) study (max = 53%) where participants were provided with an explicit list of emotions to choose from. Rather, Ojha's (2013) findings are closer to the rates we found for mismatching upfixes (Experiment 1: ~40%, Experiment 2: ~45%), supporting that neutral faces were actually mismatches. The fact that higher rates of agreement for interpretations in our study were provided by freely given responses further supports that comic readers are able to recognise the explicit meanings of upfixes when they match their faces.

We take these results to indicate that comprehenders draw on an abstract schema in the understanding of upfixes. Such a schema is posited as abstracted across observations of the conventional, item-based face–upfix pairs that become stored in long-term memory. This interpretation would imply that, the more experience a comprehender has with upfixes and comics (where upfixes are mostly found), the more they will be able to generalise an abstract visual morphological class. Such an interpretation is at least suggested by an early study by Newton (1985), who found that children's understanding of upfixes was modulated by both age and the frequency that those upfixes appeared in comics. We would thus extend such findings to hypothesise that viewing multiple *types* of upfixes allows for generalisation across these learned pairs to form an abstract schema.

In line with this, our correlations with VLFI scores suggested that more experience reading comics modulates this recognition between matching and mismatching face–upfix pairs. In Experiment 1, greater comic reading experience correlated with larger differences between familiarity ratings of conventional and unconventional mismatches but smaller differences between conventional matches and mismatches. In Experiment 2, greater comic reading expertise correlated with larger differences in the interpretations between conventional displaced and displaced mismatching face–upfix pairs, and smaller differences between unconventional displaced and mismatching face–upfix pairs. Finally, larger VLFI scores also implicated an effect of conventionality, with a greater difference between the agreement of interpretations of conventional and unconventional displaced upfixes.

These results suggest that knowledge of these particular upfixes, and the generalisation across them of an abstract upfix schema, is acquired as part of a "fluency" in the visual vocabulary used in comics (Cohn, 2013). Such findings are consistent with previous work showing that understanding of visual morphology increases with age and/or frequency reading comics (Cohn & Maher, 2015; Friedman & Stevenson, 1975; Nakazawa, 1998, 2004, 2005). These findings are also commensurate with the expectation that greater experience would lead to increased sensitivity to violations of these constraints (Cohn, 2013; Forceville, 2011; Newton, 1985). Previous work has shown that VLFI scores modulate neural responses suggestive of incongruity to violated (reversed) motion lines in visual narratives (Cohn & Maher, 2015). Whether similar results would maintain for the processing of upfixes would be important for future research, both within and outside of narrative contexts.

These results support that upfixes are abstract schemas subject to particular constraints, and are not merely memorised on an item-specific basis. Yet, it remains an open question whether they constitute a unique case or whether similar combinatorial constraints operate on other visual morphology, either within or outside visual narratives. In previous work, we have argued that several elements of the visual vocabulary used in comics involve abstract schema and/or morphological processes analogous to those in verbal morphology such as affixation, substitution/suppletion, and reduplication (Cohn, 2013). Would other schema be restricted by comparable constraints, or are upfixes an isolated case? While some work has suggested constraints operating on word balloons (Forceville, Veale, & Feyaerts, 2010) and motion lines (Cohn & Maher, 2015; Ito, Seno, & Yamanaka, 2010), further study would be required for various visual vocabulary items, both within and across cultures' unique graphic conventions.

Finally, this experiment has not addressed the specific cognitive processes guiding these combinations. Given the broad analogy between verbal and visual morphology, it is worth asking whether combinatorial principles across domains engage similar underlying cognitive resources, or whether these constraints require domain-specific processing. While this analogy between the "morphology" of verbal and visual languages does not mandate shared cognitive mechanisms (Cohn, 2013), similar neurocognitive responses are evoked by violations to the "grammar" of sequential images as by violations of syntactic structure in sentences (Cohn et al., 2014; Cohn et al., 2012; Sitnikova et al., 2008). Thus, it is not inconceivable to posit that combinatorial rules used to construct the units *within* such sequences—the morphology of words or images—may also recruit similar cognitive processing. Indeed, neurocognitive responses similar to those shown to language have appeared to combinatorial aspects of motion lines in visual narratives (Cohn & Maher, 2015) and natural scenes (Võ & Wolfe, 2013), which has already suggested the potential for such overlap.

Altogether, these findings provide initial support for combinatorial principles underlying the comprehension of visual morphology. These results suggest that the construction of meaning in the graphic form —at least in the structure originating in comics— uses complexity beyond recognising individual visual signs. Rather, "fluent" readers may generalise across conventional items to derive novel meanings from an abstract schema stored in memory for graphic meanings above the head.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bateman, J. A., & Wildfeuer, J. (2014). Defining units of analysis for the systematic analysis of comics: A discourse-based approach. *Studies in Comics*, *5*(2), 373–403. doi:10.1386/stic.5.2.373_1

Booij, G. (2010). *Construction morphology*. Oxford: Oxford University Press.

Cohn, N. (2007). A visual lexicon. *Public Journal of Semiotics*, *1*(1), 53–84.

Cohn, N. (2010). Extra! extra! semantics in comics!: The conceptual structure of *Chicago Tribune* advertisements. *Journal of Pragmatics*, *42*(11), 3138–3146. doi:10.1016/j.pragma.2010.04.016

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.

Cohn, N. (2016). Linguistic relativity and conceptual permeability in visual narratives: New distinctions in the relationship between language(s) and thought. In N. Cohn (Ed.), *The visual narrative reader* (pp. 315–340). London: Bloomsbury.

Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, *64*, 63–70. doi:10.1016/j.neuropsychologia.2014.09.018

Cohn, N., & Maher, S. (2015). The notion of the motion: The neurocognition of motion lines in visual narratives. *Brain Research*, *1601*, 73–84. doi:10.1016/j.brainres.2015.01.018

Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2012). (Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension. *Cognitive Psychology*, *65*(1), 1–38. doi:10.1016/j.cogpsych.2012.01.003

Cohn, N., & Wittenberg, E. (2015). Action starring narratives and events: Structure and inference in visual narrative comprehension. *Journal of Cognitive Psychology*, *27*(7), 812–828. doi:10.1080/20445911.2015.1051535

Engelhardt, Y. (2002). *The language of graphics* (Doctoral dissertation). University of Amsterdam, Amsterdam.

Feng, D., & O'Halloran, K. L. (2012). Representing emotive meaning in visual images: A social semiotic approach. *Journal of Pragmatics*, *44*(14), 2067–2084. doi:http://dx.doi.org/10.1016/j.pragma.2012.10.003

Forceville, C. (2005). Visual representations of the idealized cognitive model of anger in the asterix album *La zizanie*. *Journal of Pragmatics*, *37*(1), 69–88.

Forceville, C. (2011). Pictorial runes in *tintin and the picaros*. *Journal of Pragmatics*, *43*, 875–890.

Forceville, C., El Refaie, E., & Meesters, G. (2014). Stylistics in comics. In M. Burke (Ed.), *The routledge handbook of stylistics* (pp. 485–499). London: Routledge.

Forceville, C., & Renckens, T. (2013). The GOOD IS LIGHT and BAD IS DARK metaphor in feature films. *Metaphor and the Social World*, *3*(2), 160–179.

Forceville, C., Veale, T., & Feyaerts, K. (2010). Balloonics: The visuals of balloons in comics. In J. Goggin & D. Hassler-Forest (Eds.), *The rise and reason of comics and graphic literature: Critical essays on the form* (pp. 56–73). Jefferson: McFarland & Company.

Friedman, S. L., & Stevenson, M. B. (1975). Developmental changes in the understanding of implied motion in Two-dimensional pictures. *Child Development*, *46*, 773–778.

Ito, H., Seno, T., & Yamanaka, M. (2010). Motion impressions enhanced by converging motion lines. *Perception*, *39* (11), 1555–1561.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Kennedy, J. M. (1982). Metaphor in pictures. *Perception*, *11* (5), 589–605.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

McCloud, S. (1993). *Understanding comics: The invisible Art*. New York, NY: Harper Collins.

Nakazawa, J. (1998). *Development of manga notation understanding* (p. 182). Proceedings of the 9th Annual Conference of Japan Society of Developmental Psychology.

Nakazawa, J. (2004). Manga (comic) literacy skills as determinant factors of manga story comprehension. *Manga Studies*, *5*, 7–25.

Nakazawa, J. (2005). Development of manga (comic book) literacy in children. In D. W. Shwalb, J. Nakazawa, & B. J. Shwalb (Eds.), *Applied developmental psychology: Theory, practice, and research from Japan* (pp. 23–42). Greenwich, CT: Information Age.

Newton, D. P. (1985). Children's perception of pictorial metaphor. *Educational Psychology*, *5*(2), 179–185. doi:10.1080/0144341850050207

Ojha, A. (2013). *An experimental study on visual metaphor* (Doctoral dissertation). International Institute of Information Technology, Hyderabad, India.

Shinohara, K., & Matsunaka, Y. (2009). Pictorial metaphors of emotion in Japanese comics. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 265–293). New York: Mouton De Gruyter.

Sitnikova, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, *20*(11), 2037–2057.

Slepian, M. L., Weisbuch, M., Rutchick, A. M., Newman, L. S., & Ambady, N. (2010). Shedding light on insight: Priming bright ideas. *Journal of Experimental Social Psychology*, *46*, 696–700.

Võ, M. L.-H., & Wolfe, J. M. (2013). Differential electrophysiological signatures of semantic and syntactic scene processing. *Psychological Science*, *24*(9), 1816–1823. doi:10.1177/0956797613476955

Walker, M. (1980). *The lexicon of comicana*. Port Chester, NY: Comicana.