Combinatorial morphology in visual languages

Neil Cohn
Tilburg University
neilcohn@visuallanguagelab.com

To appear as:

Cohn, Neil. Combinatorial morphology in visual languages. In Booij, Geert (Ed.). *Advances in Construction Morphology*. London: Springer.

Abstract Just as structured mappings between phonology and meaning make up the lexicons of spoken languages, structured mappings between graphics and meaning comprise lexical items in visual languages. Such representations may also involve combinatorial meanings that arise from affixing, substituting, or reduplicating bound and self-standing visual morphemes. For example, hearts may float above a head or substitute for eyes to show a person in love, or gears may spin above a head to convey that they are thinking. Here, we explore the ways that such combinatorial morphology operates in visual languages by focusing on the balance of intrinsic and distributional construction of meaning, the variation in semantic reference and productivity, and the empirical work investigating their cross-cultural variation, processing, and acquisition. Altogether, this work draws these parallels between the visual and verbal domains that can hopefully inspire future work on visual languages within the linguistic sciences.

Keywords
Visual language; drawings; visual morphology; metaphor; combinatorial structure.

1 Introduction

If drawings only conveyed things as they looked in the world, many graphics would appear downright bizarre. Lines trailing a moving object denote movement, and gears above a head no longer represent just a machine, but thinking. Meanwhile, hearts or dollar signs may substitute for someone's eyes to convey lust or desire for money. Conventional patterns like these require the construction of meaning beyond just iconic perception, and indeed, such forms have been used in drawing systems for thousands of years (Wichmann and Nielsen 2016; Díaz Vera 2013b, 2013a; Petersen 2011). Because of their combinatorial qualities, these forms have frequently been compared to lexical items in language (Cohn 2013b; Forceville 2011; McCloud 1993; Walker 1980). We here explore this linguistic and combinatorial nature.

2 Visual Language Theory

Comparisons between graphic communication and language have recently been formalized in ***Visual Language Theory*** (VLT), which argues for parallels between the structure and cognition of language and drawing (Cohn 2013b). A language consists of a system of patterns in the mind/brain of a speaker. To the extent that the patterns of a person's idiolect are similar to those in other people, they share a common language. This basic principle can be applied not only to verbal or signed languages, but also to the representations and mechanisms used in producing and comprehending drawn, graphic information. A "language" is thus a system of expression shared across a population using a modality (phonology, graphics) mapped to meanings to create lexical items (words, images), which are ordered using a sequential grammatical system (syntax, narrative) (Cohn 2013b). VLT thus proposes that similar structures and mechanisms in the mind/brain extend across domains, with variation between systems arising from differences motivated by the modalities themselves. Thus, shared systematic sequential meaningful sounds constitute spoken languages of the world, while shared structured systematic sequential images are manifested by ***visual languages***. VLT thus draws parallels between the structure and cognition of verbal and signed modalities and that of the visual-graphic modality, and incorporates all three systems into a single cognitive architecture that allows for multimodal interactions (Cohn 2016b).

Visual languages arise in many socio-cultural contexts —instruction manuals, art, aboriginal sand drawing, emoji, etc.— and especially in the highly consistent and codified systems used in comics. Just as spoken languages vary between populations, visual languages differ based on different cultural and functional contexts. For example, mainstream superhero comics use a particular dialect of American Visual Language, which contrasts with the Japanese Visual Language stereotypically used in manga. In addition, many visual languages extend beyond one context; the conventions used in comics also appear in emoji and other aspects of visual culture (Forceville et al. 2014).

We here focus on this visual vocabulary: the structured mappings between form (graphics) and meaning that make up drawings. In particular, this article examines the combinatorial qualities of morphology that extend beyond iconic representations. This discussion will attempt to illustrate that structures of visual lexical items parallel those found in other linguistic systems, and that these structures can be formalized using the methods of the linguistic sciences.

3 Visual morphology

Like verbal languages, the lexical items in a visual language can be categorized as either ***open-class*** or ***closed-class*** (Cohn 2013b). Open-class lexical items easily allow for new patterns to be created. In visual form, these are typically iconic representations: it is easy to create a novel schematic pattern for iconic elements, based on the way they look (such as schemas for people, animals, plants, buildings, and their subcomponents, etc.). Closed-class lexical items require more conventionalization and are thus more constrained in creating novel forms. These are typically elements that depict invisible or non-iconic elements, such as motion lines (Figure 1b), speech balloons (Figure 1a), or hearts above the head to mean love (Figure 2a). These latter elements are the most interesting in terms of combinatoriality because they exhibit the most similarities to languages in other domains.

Combinatorial qualities arise in the lexicons of visual languages between forms that can stand alone (like an image of a person), and those that cannot (like a motion line). The latter elements must attach to another element, and thus have been likened to *bound morphemes* in other linguistic systems (Cohn 2007, 2013b; Engelhardt 2002; Forceville 2011).

Visual languages use similar basic strategies to combine morphemes as verbal languages (Cohn 2013b). For example, affixation in verbal languages attaches an affix to a stem in front (prefix; *un*happy), behind (suffix; jump*ed*), inside (infix; abso-*frickin'*-lutely), or around (circumfix; *en*light*en*). Similarly, visual elements attach in "upfixes" (affixes that are "up") such as lightbulbs, hearts, stars, etc. which appear *above* a character's head. Attachment also occurs between the affixes of word balloons to a speaker (Figure 1a), motion lines to a mover (Figure 1b), impact stars to a collider (Figure 1c), or anxious sweat drops to a worrier (Figure 1d). None of these affixes could stand alone if their stems were omitted.

Visual languages can also use substitution. In verbal language, substitution appears in suppletion (*go* → *went*) and internally for umlaut (s*i*ng → s*a*ng). Whole unit substitution occurs visually when a figure spins to become a tornado (Figure 1e), when several people fight inside a "fight cloud" with arms and legs sticking out (Figure 1f), or when something becomes invisible with dotted outlines (Figure 1g). Internal substitution occurs for "eye-umlauts" where the eyes of a character are replaced by hearts (lust; Figure 1h), stars (desire for fame), dollar signs (desire for money), etc. A character-specific suppletion happens to Spider-Man when he detects danger: half of his ordinary face is depicted with his mask, even though he is not wearing the costume.

Finally, visual languages also might repeat forms, just as verbal languages use reduplication (*salad-salad*; *tick-tock*). Repetition occurs overtly when all or part of a figure is shown in different postures to indicate movement (Figure 1i). It also occurs when lines repeat with slight offset to depict shaking, or alternatively the double-vision of another character (Figure 1j).

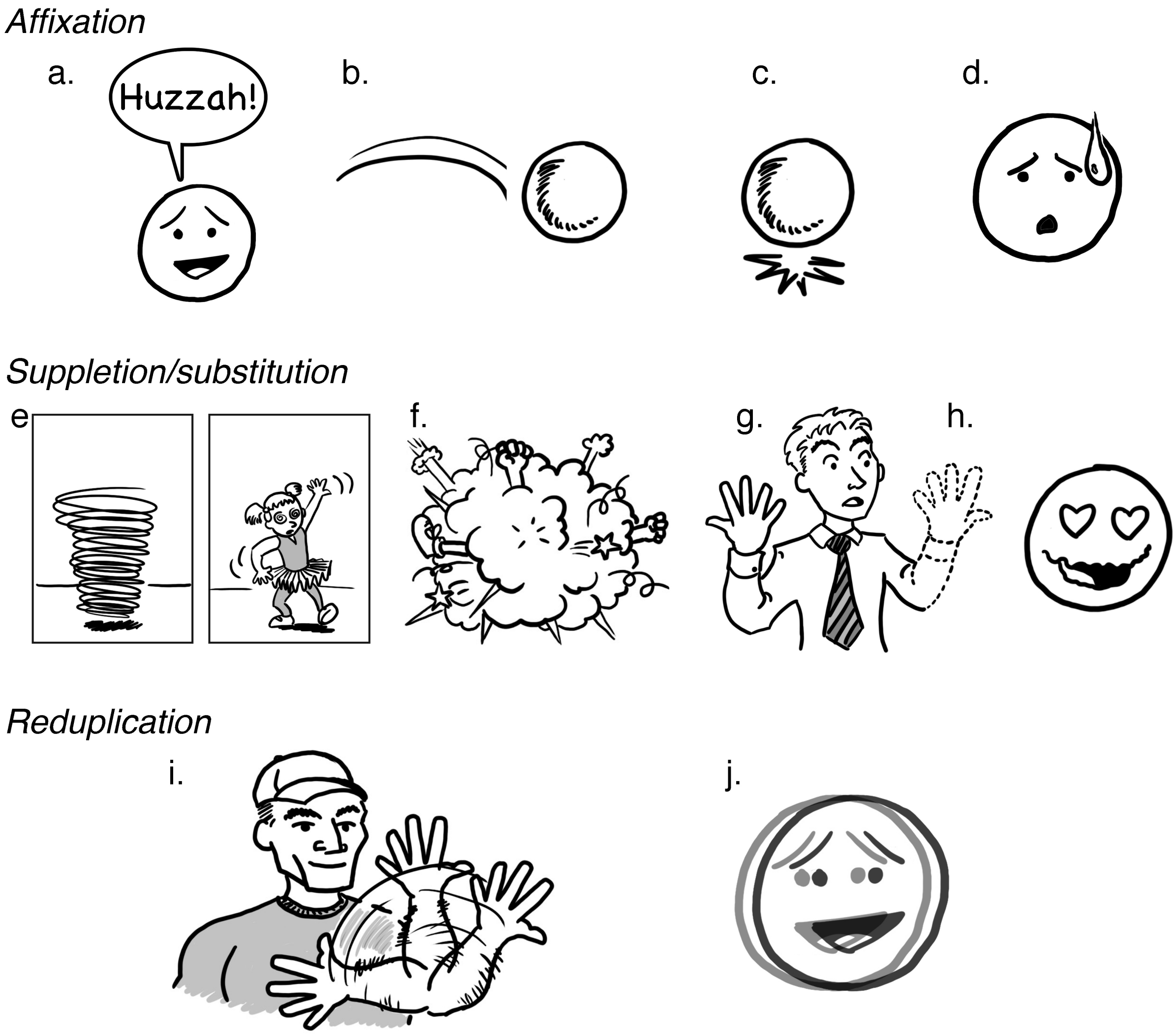


Figure 1. Visual morphological relations using a. affixation, b. suppletion/substitution, or c. reduplication.

Thus, similar basic strategies of morphological combination occur in the verbal and visual modalities. Note that these options comprise virtually all possible relations of forms: attachment, insertion, substitution, repetition, etc. This similarity does not necessarily mean that visual languages "are like" verbal languages in their morphological combinations, but rather more likely that all modalities—visual, verbal, and signed—make use of available combinatorial possibilities.

4 Units of visual morphology

If bound morphemes attach to stems in visual forms, what are the "base units" and "morphemes"? This question bears on a persistent critique of applying linguistic theories to graphic information: Just what is a minimal unit? I will argue that searching for a "minimal unit" is misguided. First, the visual modality deals with meaningful information in different ways than the verbal domain. The analog nature of visual information allows for graphic representation to embed meaningful (and schematized) forms in each other simultaneously, rather than in a temporally sequential fashion as in speech. Thus, the differences in modality are not parallel in this regard. Second, the language sciences have acknowledged that structuralism's focus on minimal units is no longer tenable. Rather, within linguistics itself there is a strong and persistent critique of the notion of minimal meaningful units in morphology (Booij 2010; Jackendoff 2002; Jackendoff and Audring 2016; Sadock 1991; Anderson 1992; Stump 2001), where the emphasis in some circles is on the mapping between form and meaning —here, the form being graphic rather than phonological— regardless of the size of the lexical item.

Graphic forms do differentiate between elements that can stand alone and those that cannot, just as words can stand alone, but bound morphemes cannot. For example, though a head is recognized as necessarily attaching to a body, it has more ability to stand alone as a perceptual form than an isolated nose, or even just a body (such as when a body is depicted in a comic panel, but its head lies outside the frame).

We can thus introduce a notion for a visual isolatable form—whether or not it may also be divisible into smaller morphemes. This is roughly analogous to the level of a "word" in verbal language. I will call this form a ***monomorph***.[1] Similar to the morphological level of a "word," a monomorph is *a visual representation that can stand alone without needing to attach to other morphemes, can be made up of smaller morphemes, and can combine with other elements to form even larger monomorphs*.

Because monomorphs are perceptual objects that can stand alone as an isolatable "morphological bundle," we should be able to crop these elements (as with the frame of a comic panel) and still recognize them as the same entity. For example, a character's face, bust, or whole body would all be sufficient for identifying the identity of that character. For animate entities, the face may thus be "marked" as an essential identifying feature of monomorphs. Cropping that limits identification and would depict less than a monomorph (i.e., eyes alone, a hand, body, etc.), though would still be recognizable *as part of* a potentially isolatable object. These "morphological bundles" that are subsets of a monomorph will be termed as ***micromorphs***. The quantity of information does not matter—as long as it falls beneath the level of a monomorph it classifies as a micromorph. Thus, as in other part-whole relations, depiction of just part of an entity (a micromorph) would entail the existence of the whole (monomorph).

It is important to stress that this theory of morphology does not aim to describe meanings (i.e., semantics). Rather, this level of analysis describes the relations between *forms* alone, and these forms interface with semantics (in a manner formalized below). In verbal language, a word like *disbelieve* consists of two morphological units: a word (*believe*), which can stand alone, and an affix (*dis-*) which cannot stand alone, and must attach to another object. This morphological information then maps to the semantic information of NOT-BELIEVE. This theory of visual morphology posits an analogous relationship, as it aims to describe the visual forms that map to semantic information.

---

[1] A precedent for this exists in Koch's (1971) proposal for "logemes" which are made up of visual morphemes, though Koch did not distinguish "stems" versus bound morphemes. For a summary in English, see Wildfeuer and Bateman (2016).

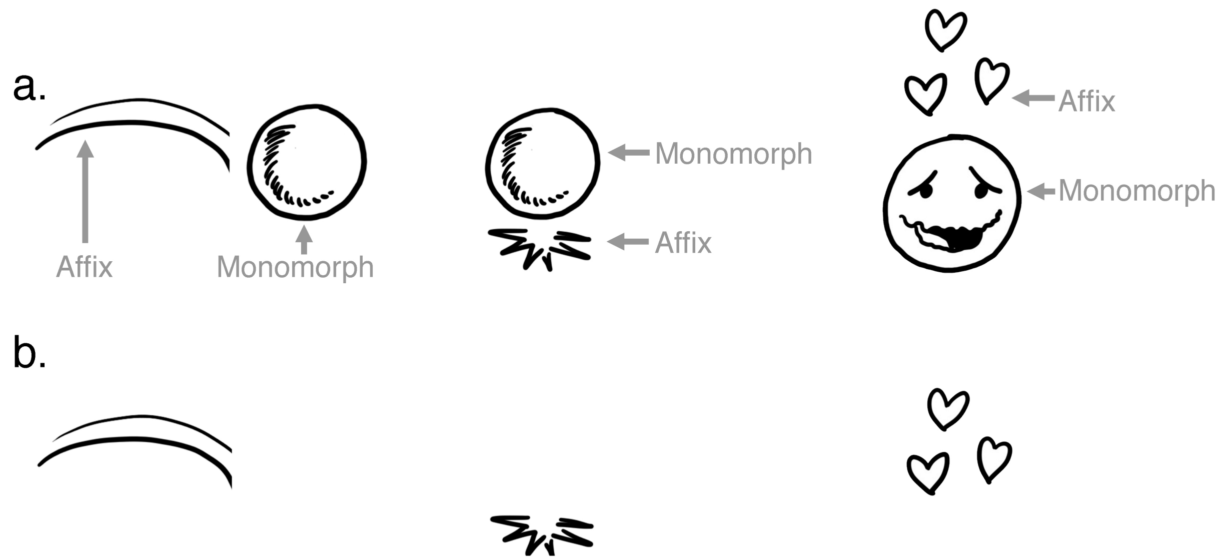


Figure 2. Examples of affixes attaching to monomorphs (moving ball, impact star, hearts above head).

The bound morphemes in the visual modality thus attach to a ***stem*** of a monomorph to form a larger composition. For example, motion lines attach to a moving object, impact stars to an impacting object, and upfixes to a face to depict an emotional/mental state. As depicted in Figure 2a, each of these affixes attaches to a monomorph. These are all bound morphemes, evident through the odd and incomplete nature when their monomorphs are deleted, as in Figure 2b. This bound and compositional nature can thus be captured in a general schema as:

(1) $[_{\text{Monomorph}}$ {Monomorph – (Affix)}]

It is important to stress that, in line with constructional models of morphology (Booij 2010; Jackendoff and Audring 2016), this expression is a schema stored in memory, not a procedural rule. The schematic notation here articulates that a monomorph is made up of a monomorph and an (optional) affix. Also, order does not matter (notated by the curly brackets). Since the graphic structure may specify different physical relations, it is unimportant whether our notation reflects spatial relations (i.e., for an upfix, we do not need to specify in this notation that the "affix" is above the "monomorph"). In most cases, monomorphs are stems, though some affixes can attach to micromorphs (such as a cropped image of a hand with stars indicating pain).

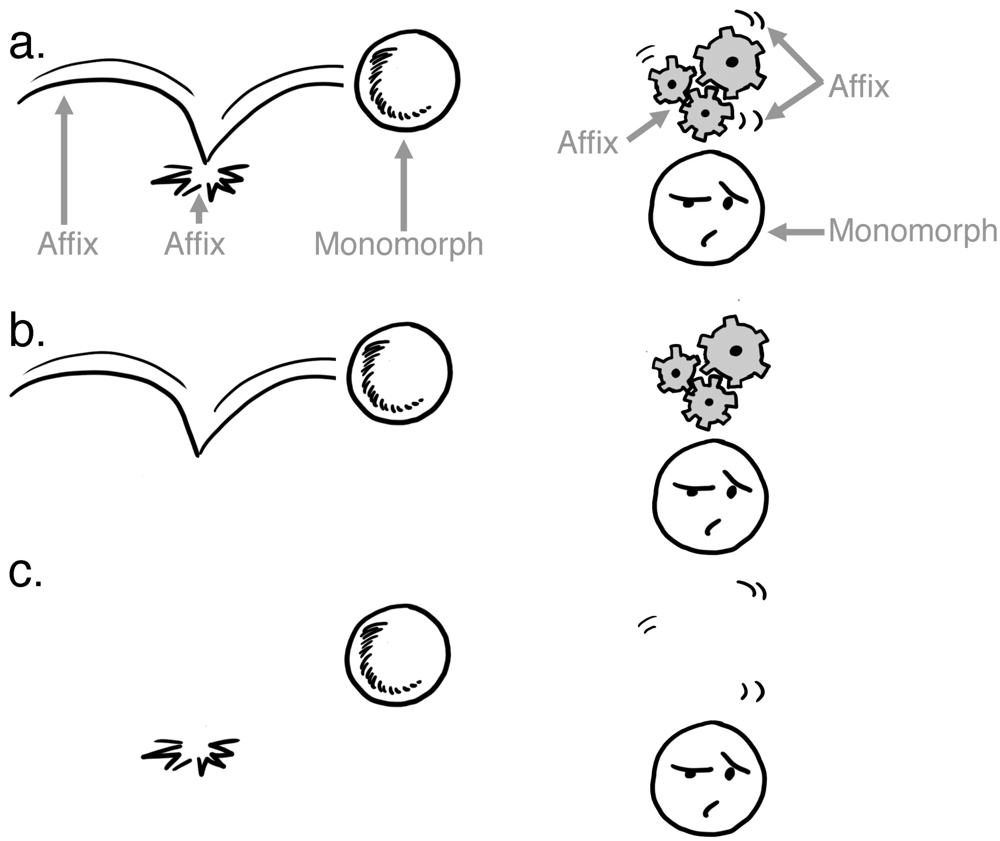


Figure 3. Affixes that attach to other affixes, and deletion tests to illustrate these dependencies.

In addition, this process can be hierarchical: affixes can serve as the stem for other affixes. Consider now the morphology in Figure 3a. Here, the impact star does not attach to a monomorph, but to the affix of a motion line. In addition, the "circumfixing" motion lines surround the affix of the gears, not the monomorph of the head. Deletion tests further show that these elements do not affix to their stems, but rather to the other affixes. As shown in Figure 3b, omission of these affixes leaves a coherent representation (though it changes the meaning). However, omission of the primary affix renders the image incoherent (Figure 3c), as would omission of the stem. This suggests that in the visual form, affixes can possibly combine with other affixes. We can generalize this as:

(2) $[_{\text{Affix}}\{\textit{Affix} - (\text{Affix})\}]$

This schema outlines that a primary, "head" affix (italics) can attach to another affix within a larger affix. Often a clear hierarchic relationship exists between affixes: the impact star attaches *to* the motion lines, not both equally applying to their stem (again, see Figure 3c).

We have thus arrived at basic constructs in visual morphology, which involves mappings of form-meaning information. Clusters of morphological information that can stand alone are monomorphs, while sub-sections of monomorphs that cannot stand alone are micromorphs. Affixes are bound morphemes that must attach to monomorphs (and sometimes micromorphs), and in certain cases can also attach to other affixes.

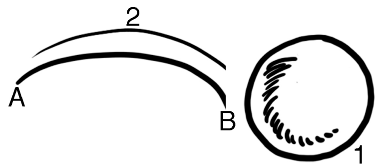
5 The parallel architecture

Having defined basic aspects of visual morphology, as separate from semantics, we can now incorporate this information into a larger model that specifies these relationships. Visual Language Theory is embedded within the broader linguistic framework of the ***parallel architecture*** (Jackendoff 2002; Jackendoff and Audring 2016), which balances different forms of structure that make up a linguistic system (Cohn 2016b). For example, a word (like *cat*) reflects the linkages between phonology (/kæt/), morphology (word), syntax (noun), and semantics (CAT) that are encoded in memory as a lexical item. Similarly, visual morphology involves an interaction between a graphic structure (the lines and shapes that make up a depicted image), a morphological structure (discussed above), and a conceptual structure of semantics.

Consider the affix of a motion line in Figure 4, which depicts a ball moving. Beneath the image is a notation for the "lexical entry" of the motion line in combination with an object. Each structure describes a separate type of information involved in the representation, which are cross-listed using indices. The graphic structure (GS) should specify the graphic elements involved in depicting these visual forms as lines, curves, dots, and the rules of their combinations. This requires a theory with the sophistication of phonology that would be beyond the scope of this chapter. Here, I instead characterize the *relational* aspects of graphic elements. These relations involve operations juxtaposing elements, substituting one for another, inserting one into another, fusing them together, distorting them, etc. (Phillips and McQuarrie 2004; Schilperoord 2013, 2017), similar to the morphological strategies described above (affixation, suppletion, reduplication). It is acknowledged that these variables do not describe purely graphic structures, but rather interactions between graphics and morpho-semantics. For our purposes of describing how morphological information manifests visually though, they will suffice.

In the motion lines of Figure 4, the graphic relations here *juxtapose* the motion lines (subscript "2") with the objects they affix to (subscript "1"). This juxtaposition is fairly content neutral—it does not specify whether the lines are in front or behind the object, only that one graphic shape is next to another one. The contents of those shapes are determined by the semantics (see below). This meaning also generates an inference that motion lines are placed behind their object, which prevents, for example, the incongruous positioning of motion lines ahead of their stem (Cohn and Maher 2015; Ito et al. 2010).

The morphological structure (MS) specifies that motion lines are an affix (1), which attaches to a monomorph (the ball: 2) to form a larger monomorph (subscript "i"). Finally, the conceptual structure (CS) specifies the meaning of a motion line, here notated using Jackendoff's (1990) Conceptual Semantics. It says that an object X (here, the ball) goes along a path from one place (A) to another (B), with the motion line depicting the path (subscript 1). Because of this path information, we therefore infer that the graphic element of the motion line falls *behind* its juxtaposed object. Together, these pieces of structure combine to give the overall structural understanding of a basic motion line.



Motion Line
GS: [JUXTAPOSE(2,1)]
MS: $[_{Mm}$ Monomorph$_1$ - Affix$_2]_i$
CS: $[_{Event}$ GO($[_{Object}$ X$_1]$;$[_{Path}$ FROM(A), TO(B)$]_2$)$]_i$

Figure 4. A lexical entry for motion lines within a parallel architecture.

This type of formalization is useful for two reasons. First, formalizing these structures allows us to be specific about their component elements. Such specificity also allows for predictions, which can be examined with experimentation (see below). Second, formalization allows us to represent the componential parts of a visual lexicon in line with lexical entries for other modalities of language. The parallel architecture thus predicts that the cognitive instantiation of visual morphemes encodes links between different domains of structure in memory—and the involved structures may operate across modalities (for example, conceptual structure specifies the meaning of both verbal and visual forms).

6 Complementary distributionality

As in other languages, the meaning of visual morphemes arises from an interaction between the signs themselves and their contextual distribution (Cohn 2013b; Forceville 2011; McCloud 1993). In visual morphology, part of that context is a location and orientation relative to other graphic elements (Cohn 2013b; Forceville 2011). This section will explore these relationships.

Let's begin with the morpheme of a heart, which retains largely the same meaning when placed in different locations. Figure 5a formalizes the structure of hearts as upfixes. First, the graphic structure places the upfix (here hearts) as juxtaposed with the head/face. In this case, the juxtaposition is specific, placing the upfix *above* the head/face (Cohn 2013b; Cohn et al. 2016), and that upfix is oriented vertically (Forceville 2011). The morphological structure again simply links the monomorph (head/face) to the affix (upfix) to form a larger monomorph.

Finally, the conceptual structure is depicted in two parts. First, the box below describes the "components" of the upfix: this is the meaning for the face and hearts independently.[2] The composite meaning appears in the main lexical entry (i.e., the meaning that results from combining those component parts). On its own, the face shows that the person (X) is in a state of being happy (1,i). The hearts alone mean love (2,j). Their union means that the person is happily

---

[2] If being complete, we could specify full lexical entries (GS, MS, CS) for the face, the upfix, and their combination separately. For concision, I here just notate the combinatorial whole in full, and the componential aspects of the conceptual structure.

in love (i,j), as in the resultant meaning listed in the lexical entry. This compositional meaning is fairly straightforward.

Now let's consider the "eye-umlaut" using hearts, where they replace eyes (Figure 5b). This overall meaning is largely the same as the upfix. In this case, the graphic structure is slightly more complex, because it uses substitution/suppletion (hearts into eyes) rather than affixation (hearts juxtaposed above head). In order to substitute hearts for eyes, it requires comparing the eye-umlaut face (1) to a default, basic face in memory that has eyes intact (1'). All aspects of the eye-umlaut face are thus co-indexed to the default face through paradigmatic relationships. The contrast between these faces uses a relation of "same-except" (Culicover and Jackendoff 2012) whereby some of the *same* features appear in both, *except* the crucial components of interest. Here, those exceptions are the substitution of the hearts for the eyes. This is formalized with a function of SAME applying both to the manifested face (1) and the basic schematic face (1') to notate the similarities between the faces, and with the function of EXCEPT specifying the substitution of eye-umlauts (2) for eyes (2').

The morphological relationship between heart eye-umlauts and heart upfixes does not differ. Both use a monomorph (face) and affix (hearts), but in different graphic relations. In addition, the semantics also subtly differ. The componential parts are the same as the upfix (face, hearts), but the compositional meaning has a slight variation. Because the hearts substitute for eyes, they do not just mean that the person is in the state of being in love (as with upfixes), but that the person sees in a loving/lustful manner. The substituted element (eyes) thus contributes to the compositional semantics.

a.

Heart Upfix

GS: $[\text{ABOVE}([_{\text{Upfix}}\ Z_2],\text{HEAD}_1)]$

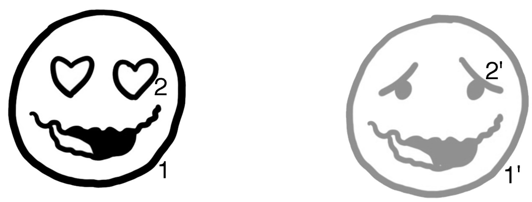
MS: $[_{\text{Mm}}\ \text{Monomorph}_1 \text{ - Affix}_2]$

CS: $[_{\text{State}}\ \text{BE}([_{\text{Object}}\ X_1],[_{\text{Property}}\ \text{HAPPILY IN-LOVE}])]_{ij}$

↕

Components CS:
Face: $[_{\text{State}}\ \text{BE}([_{\text{Object}}\ X_1],[_{\text{Property}}\ \text{HAPPY}])]_i$
Hearts: $[_{\text{State}}\ \text{LOVE}(X,Y)]_{2,j}$

b.



Heart Eye-Umlauts

GS: $[\text{SAME}(\text{FACE}_1,\text{FACE}_{1'});\text{EXCEPT}(\text{SUBSTITUTE}([_{\text{Eye-Umlaut}}\ Z_2],\text{EYES}_{2'}))]$

MS: $[_{\text{Mm}}\ \text{Monomorph}_1 \text{ - Affix}_2]$

CS: $[_{\text{Event}}\ \text{SEE}([_{\text{Object}}\ X_1],[_{\text{Property}}\ \text{LUSTFULLY}_2])]_{ij}$

↕

Components CS:
Face: $[_{\text{State}}\ \text{BE}([_{\text{Object}}\ X_1],[_{\text{Property}}\ \text{HAPPY}])]_i$
Hearts: $[_{\text{State}}\ \text{LOVE}(X,Y)]_{2,j}$

Figure 5. The visual morpheme of hearts with largely similar meanings in different locations.

Overall, hearts use the same morpheme with largely the same meaning in different locations. This is in part facilitated by the relatively fixed semantics of the heart itself which is widespread beyond specific visual languages (Forceville et al. 2014), though different locations add nuanced change to the meaning. Now let's compare this to a star, which has totally different meanings depending on placement as an upfix, eye-umlaut, or affix.

First, stars as upfixes (Figure 6a) mean dizziness or disorientation. This upfix uses a compounding of affixes of a group of stars and curved motion lines. Stars, unlike hearts, on their own have no intrinsic meaning besides being a shape. This is notated in conceptual structure with an object (here the stars, "Y") going along a circular path repeatedly (notated with the plural marker PL). Together with the face, these morphemes give the idea that the person is dizzy, a marginally compositional meaning. Such meaning rests in part on the English idioms "seeing stars" and "head spinning" to mean dizziness, which are entrenched in memory (depicted here in grey) and invoked by the composite morphemes of stars and a circular motion line path.

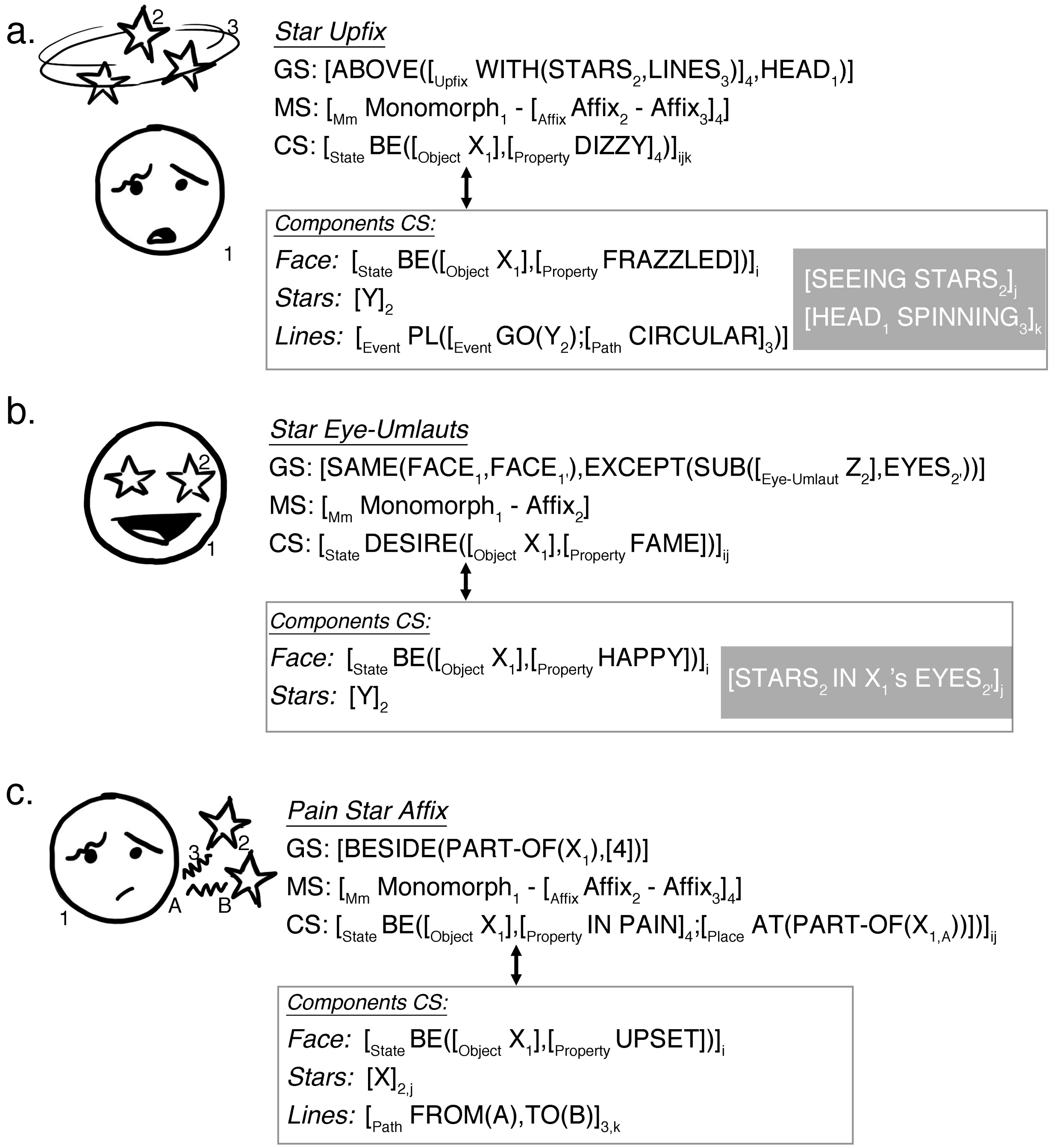


Figure 6. Stars appearing as different affixes based on location.

Now consider the eye-umlaut in Figure 6b. Here, stars substitute for eyes and have nothing to do with seeing (despite the idiom "seeing stars" invoked by an upfix) or with dizziness. Rather, they mean a desire for fame. Again, as with hearts, this eye-umlaut invokes a more basic schematic face (not depicted) with a same-except function. On its own the face is in a state of being happy, with the stars not having any meaning. These morphemes together lack any compositionality. Rather, the final meaning of desire for fame arises through their connection to an entrenched verbal idiomatic expression of “stars in one’s eyes” (grey box).

Finally, stars as affixes mean pain, as in Figure 6c. Here, stars combine with radial lines to a stem of some body part. The stars again convey no meaning on their own (2) but the radial lines specify a path for the stars to radiate from (A). This location is thus the place where the person feels pain. Some uses of pain stars lack radial lines, instead signaling a general, non-focal pain to the area they surround. In this affix, the stars use a conventionalized meaning, with no reference to an additional entrenched idiom.

Thus, in contrast to the relatively stable meanings for hearts in different locations, stars change their meanings entirely based on their distribution. Because of this, the stars themselves may be considered as different morphemes with a similar graphic structure. Similar variation occurs between verbal homonyms, such as the difference between the *un-*’s that negate adjectives (*unhappy*) and reverse actions (*untie*), or the *-er*’s that compare properties (*stronger*) and doers of actions (*baker*). Note also that one cannot “switch” the meanings between distributions. While the pain affixes may apply somewhat to the upfix (particularly if not spinning), the eye-umlaut stars cannot be seen as pain in the eyes. Similarly, the stars as affixes cannot indicate a desire for fame. Several visual morphemes have this type of distributionally defined meaning (see Cohn 2013b; Forceville 2011; McCloud 1993 for more examples).

a.

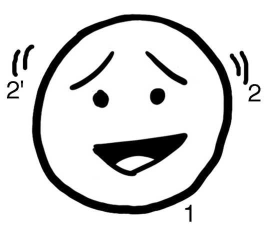
GS: [REPEAT(1); OFFSET(1,1')]

MS: $[_{Mm}$ Monomorph_1 - Monomorph_2]

CS1: $[_{Event}$ PL($[_{Event}$ GO($[_{Object}$ X_1];$[_{Path}$ {FROM/TO(1)}, {TO/FROM(1')}])])]

CS2: $[_{Event}$ SEE($[_{Object}$ Y],$[_{State}$ BE($[_{Object}$ X_1])];$[_{Manner}$ DRUNKENLY/DIZZILY])]

b.



GS: [SURROUND(2,1)]

MS: $[_{Mm}$ Monomorph_1 - Affix_2]

CS: $[_{Event}$ PL($[_{Event}$ GO($[_{Object}$ X_1];$[_{Path}$ {FROM/TO(2)}, {TO/FROM(2')}])])]

Figure 7. The same graphic structure using reduplication with two different meanings (a), and a different morpheme (motion lines) depicting one of those same meanings (b).

Let's examine two additional contrasts: a single morphological distribution with multiple meanings, and different morphemes with the same meanings. Figure 7a uses reduplication of the lines of a face. The graphic structure thus repeats the same lines (1) with a slight offset (1'). In full formalization, this offset would specify a particular distance—not close enough to seem like single thick lines or shadows, but not so far away as to appear like a separate entity. Multiple repetitions may also be possible. Both faces are thus monomorphs that combine to form a larger *compound monomorph*. There is no affixation here of adding a bound morpheme, but rather the base itself is altered.

Under one interpretation (CS1), this depicts shaking between one location (1) and another (1'). Again, the recursive repetition of this event is notated with the plural marker (PL). This same representation can have an alternative meaning though (CS2): It could also represent the "double vision" of a person (Y, not depicted) who is drunk or dizzy. Thus, this graphic structure is polysemous, although both meanings share a trait of instability (of object or viewer of object).

Note also that the first interpretation (shaking) can also be conveyed with circumfixing motion lines surrounding the object. Here, the offset reduplication does not mark the start and endpoints of the path, but rather this is conveyed by the affixed lines. Thus, here we have a single morphological representation with two possible meanings, and one of those meanings can be conveyed with a different morpheme. Indeed, circumfixing motion lines themselves are slightly polysemous: here their movement has a manner of *shaking*, but next to gears, as in Figure 3a (and Figure 9, further on), their manner is *spinning*. Again, note that both meanings cannot be maintained at the same time: it is difficult to retain a construal where Figure 7a is someone shaking who is *also* seen by another person with double vision.

Thus, to summarize, visual morphemes have a complicated relationship with their distribution (Cohn 2013b; Forceville 2011). Some morphemes retain largely the same meaning in different distributions (hearts), while others wholly depend on distribution to determine their meanings (stars). In addition, some single morphemes can convey multiple meanings (offset reduplicative lines), while the same meaning might be conveyed by multiple morphemes (reduplication, circumfixing motion lines).

7 Semantic variation in visual morphology

Compositional visual morphology uses a variety of types of semiotic reference to express meaning. As discussed, hearts have a fixed meaning of LOVE. Thus, when combined with a face, they rely on that intrinsic symbolic meaning, even when in different distributional locations (Figure 5). In contrast, stars have no intrinsic meaning and are wholly context dependent (Figure 6). These combinations are also illustrations of idiomatic expressions ("seeing stars" or "stars in their eyes"), which make them "permeable" (Cohn 2016a) in the sense that they are a conceptualization shared across multiple expressive modalities (here, verbal and visual).

Motion lines use another type of reference, as a depiction of an invisible, yet basic, cognitive conceptualization, i.e., paths (Jackendoff 1990; Talmy 2000). Such folk understanding of paths also occurs in *scopic lines* that use dotted lines to depict the vector from characters' eyes to what they are looking at (Figure 8a). Paths also occur with *radial lines*, like the straight lines that emerge from something shiny (like the sun) or the wavy lines used to depict heat or smells (like above trash or coffee) as in Figure 8b. These examples depict varying invisible paths. These are different from representations like *focal lines* which have a deictic function of drawing attention to something (Cohn 2013b; Forceville 2011), like eyes (Figure 8c). Focal lines are not paths, but belong to the broader category of *indexical lines* (Cohn 2013b). Note also that the radial lines on the gold in Figure 8b and focal lines on the eye in Figure 8c are another example of distributionally defined morphemes, like the stars discussed above. They share a common graphic structure (three converging lines) but have a different distribution and construed meaning, making them homonyms.

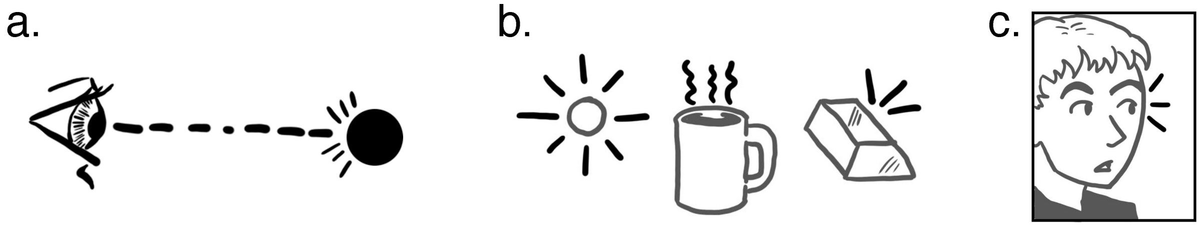
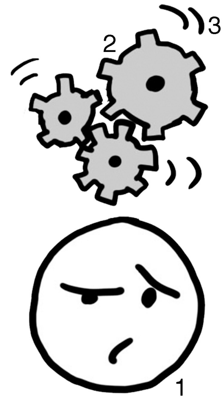


Figure 8. Path lines including a) scopic lines depicting the line of sight, b) radial lines for brightness and shine (sun, gold) and heat/smell (coffee). Also, c) depicts focal lines drawing attention to something, but without a path.

Some morphology uses iconicity. Reduplicative offset lines (Figure 7a) either show an iconic subjective point-of-view (double-vision) or attempt to replicate an iconic phenomenon (shaking). Iconic visual morphology undergoes more complexity when it is transformed by ***conceptual metaphors*** (Lakoff and Johnson 1980; Forceville 2016) whereby one semantic domain is understood in terms of another. These include spinning gears evoking the idea that the mind is a machine, or steam coming from ears as if anger is fluid in a pressurized container (elaborated on below). In a sense, these representations evoke a kind of "semiotic coercion" (e.g., Audring and Booij 2016) whereby iconic representations (gears, lightbulbs, birds) are "coerced" into a symbolic construal via their connection to an underlying metaphor and their emerging inferences.

Take for example the upfix of spinning gears to convey the meaning that the person is thinking (Figure 9). The literal semantics of this representation are of a quizzical face (1)—which requires the background knowledge that the person possesses a mind (i)—and that gears (2) are moving (3). On their own, gears have little to do with thinking. Rather, this upfix involves two entrenched conceptual metaphors (grey box): MIND IS A MACHINE and PROGRESS IS MOVEMENT (Cohn 2010; Lakoff and Johnson 1980). MIND IS A MACHINE specifies that the properties of a mind are similar to that of a machine, with gears being the mechanisms by which thinking occurs. PROGRESS IS MOVEMENT describes that advancement occurs because of motion—here, that effort is being exerted because of the motion of gears.

Together, an emergent metaphor maps moving machines (4) as being representative of *progress of the mind* (m). Moving machines are directly depicted in the image (gears with motion lines), but "progress of the mind" is an emergent, inferred property of this representation and the combination of these conceptual metaphors (thus, underlined). "Progress of the mind" is essentially "thinking" and thus yields the overall interpretation that spinning gears above the head mean thinking. Note, that both metaphors are necessary: If the motion lines were absent (as in Figure 3b), it would lose the PROGRESS IS MOVEMENT metaphor and thus the active sense of thinking as a process. Such movement means that the event is durative and ongoing (in contrast to the punctive state of inspiration with an upfix of a lightbulb). Though these complex mappings underlie the resultant meaning of this upfix, all of this information is encoded in the basic understanding of this visual lexical item.



Moving Gears Upfix

GS: [ABOVE([$_{Upfix}$ SURROUND(LINES$_3$,GEARS$_2$)]$_4$,HEAD$_1$)]

MS: [$_{Mm}$ Monomorph$_1$ - [$_{Affix}$ *Affix*$_2$ - Affix$_3$]$_4$]

CS: [$_{Event}$ THINK([$_{Object}$ X$_1$]))]$_{i,m}$

↕

Components CS:

Face: [$_{State}$ BE([$_{Object}$ X$_1$],[$_{Property}$ QUIZZICAL]))]$_i$

[$_{State}$ POS([$_{Object}$ X$_1$],[$_{Object}$ MIND$_{1'}$]))]$_i$

Lines: [$_{Event}$ MOVE([$_{Object}$ Y]$_2$)$_3$]$_4$

Gears: [$_{Object}$ PL([$_{Object}$ GEAR]$_2$)]

Entrenched metaphors:

[MIND$_1$ IS A MACHINE$_2$]$_j$

[PROGRESS$_5$ IS MOVEMENT$_3$]$_k$

Emergent metaphor:

[[$_{Event}$ PROGRESS(MIND$_{1'}$)$_5$]$_m$ IS [$_{Event}$ MOVE(MACHINE$_2$)$_3$]$_4$]$_{j,k}$

↕

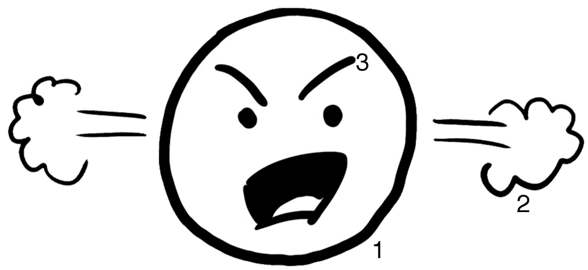
[$_{Event}$ THINK([$_{Object}$ X$_1$]))]$_{i,m}$

Figure 9. Formalized structure for gear upfix involving conceptual metaphors.

Another conceptual metaphor is involved in the representation of steam coming from a person's ears, as in Figure 10. Graphically, the steam surrounds the head on either side, meaning that steam cannot just be from one side (also note that here, we infer the ears because of the compositional meaning, despite not being depicted). This representation evokes the metaphor of ANGER IS HOT FLUID IN A PRESSURIZED CONTAINER, which appears in several visual and verbal expressions (Forceville 2005; Lakoff 1992), such as *He was steaming mad*, and also upfixes with emanating heat lines (wiggly lines above the head).

In this case, the depiction directly has a face (1) with an expression of anger (3) and steam (2) being emitted from the inferred ears. To involve steam, the entrenched metaphor must involve the understanding first that hot fluid in a pressurized container emits steam. The container is mapped to a head, and anger is mapped to the hot fluid (grey lines). But, steam remains constant in both, yielding the inference that anger in a head *emits steam* (underlined).

This then leads to the depiction of steam coming from the ears creating the overall inference that the person is *steaming angry*. Again, though this understanding uses fairly complex mappings within conceptual structure, such meaning is encoded in memory for this visual lexical item.



Steam Affix

GS: $[\text{SURROUND}(\text{HEAD}_1,[\text{STEAM;FROM(EARS)}]_2)]$

MS: $[_{Mm}\ \text{Monomorph}_1 - \text{Affix}^{x2}_2\ ]$

CS: $[_{State}\ \text{BE}([_{Object}\ X_1],[_{Property}\ \text{“STEAMING” ANGRY}]_2)]_{i,k}$

↕

Components

Face: $[_{State}\ \text{BE}([_{Object}\ X_1],[_{Property}\ \text{ANGRY}_3])]_i$

Steam emission: $[_{Event}\ \text{EMIT}([_{Object}\ X_1],[_{Object}\ \text{STEAM}_2];[_{Path}\ \text{FROM(EARS)}])]_i$

Entrenched metaphor:

$[\text{ANGER}_1 \text{ IS HOT FLUID}_4 \text{ IN A PRESSURIZED CONTAINER}_5]_j$

Emergent metaphor:

$[_{Event}\ \text{EMIT}([_{Place}\ \text{IN}([_{Object}\ \text{HOT FLUID}_4],\ [_{Object}\ \text{CONTAINER}_5])]),\ \text{STEAM}_2\ ]_j$

↕ ↕

$[_{Event}\ \text{EMIT}([_{Place}\ \text{IN}([_{Property}\ \text{ANGER}_3\ ],\ [_{Object}\ \text{HEAD}_1\ ])]),\ \underline{\textit{STEAM}_2}\ ]_k$

↕

$[_{State}\ \text{BE}([_{Object}\ X_1],[_{Property}\ \text{“STEAMING” ANGRY}]_2)]_{i,k}$

Figure 10. Formalized structure for affix of steam coming from ears involving conceptual metaphor.

Beyond metaphors, additional richness in visual representations may involve conceptual blending (Forceville 2016; Fauconnier and Turner 2002). Blending is a mapping of conceptual domains onto each other, but may not evoke an entrenched metaphor. Take for example blending that occurs in the graphic novel of *Maus* by Art Spiegelman, which chronicles his family's experiences in World War II and the Holocaust. It blends different types of people with animals: Nazis as cats and Jews as mice. These blends are overt, with the heads of the animals appearing directly on the bodies of humans. However, because of these visually depicted blends, an additional emergent inference arises—the predatory relationship between cats and mice is then mapped onto the relationship between Nazis and Jews. Blends between animals and people (typically animal heads with quasi-human bodies) are well entrenched as patterns in drawings and cartoons, and the practice extends far back in history in several cultures (Petersen 2011; Schodt 1983). However, these particular relationships (Nazis to cats, Jews to mice) do not necessarily represent conventionalized metaphors that are entrenched in memory, but rather are novel relations from Spiegelman's creativity. Because blending results from novel mappings, they do not often involve conventionalized morphology.

8. Productivity

Like the verbal and signed modalities, visual morphology makes use of several methods for creating novel forms. One method is through borrowing from other modalities directly. For instance, they may visually "translate" idiomatic expressions from speech such as stars as upfixes ("seeing stars") or eye-umlauts ("stars in his eyes") as in Figure 6. This is also found in a trail of daggers to show the path of an eye-line as in "starring daggers" at someone (McCloud 1993). Idioms may be one way that new closed-class morphemes may emerge. Similarly, depictions involving conceptual metaphor, and blending may allow for creative morphology that grows to become conventional, through systemization from repeated use.

Other lexical items may grow from extending and systematizing iconic representations. For example, Japanese Visual Language (JVL) in manga uses a tubular X-shape to depict popped-out veins for anger, originally placed iconically on characters' foreheads. Over time, this depiction became more schematized, and extended as a symbolic affix placed on various body parts (foreheads, hands, etc.) and even floating in speech balloons (Shinohara and Matsunaka 2009). A similar trajectory is suggested by a novel affix in Tatsuya Ishida's *Sinfest* (www.sinfest.net, January 5, 2017), which combined floating pain stars, as in Figure 6c, with floating X-shaped band-aids. Band-aids may not yet be conventionalized closed-class morphology (though they may be conventionalized open-class items), but their extension as affixes over repeated usage could become regularized. Thus, lexical items can grow from an unconventional iconic representation, to a systematic sign (Garrod et al. 2007), to a lexicalized affix. Many examples of closed-class visual morphology followed a similar trajectory (Petersen 2011), which is akin to the development of some lexical items in other modalities, such as sign language (Fay et al. 2014).

Existing schemas may also incorporate novel morphemes. For example, upfixes appear to be a ***semi-productive*** morphological class, whereby new forms can be generated that stay within the constraints of the pattern. That is, new items can be put above a character's head to become novel upfixes. Though they are less comprehensible than conventional upfixes, experimentation has shown that novel upfixes are rated as more comprehensible than incongruous ones (mismatches between face and upfix), and are subject to the same constraints as conventional ones. For example the upfix must be above, and not beside, the head, and the facial expression must "agree" with the upfix. For instance, storm clouds cannot be above a smiling face (Cohn et al. 2016). These results imply that upfixes use a productive schema, and are not simply memorized instances.

Such productivity may also be possible with eye-umlauts, though the signs must be small enough without detail to substitute for eyes. For example, it may be easier to create novel eye-umlauts of simple windows as eyes (for the idiom "eyes are the windows to the soul") than for a fully detailed bedroom with bed, wardrobe, and dresser (for the idiom "bedroom eyes"). The small graphic space for eyes will likely constrain the ability to convey complex visual details.

Semi-productivity may modulate the manner of other signs. For example, textures of "carriers" of text like "speech balloons" can modulate the shape or texture to convey the volume (e.g., jagged lines for yelling or dotted lines for whispering), pragmatic intent (e.g., a drippy carrier for sarcasm), or origin (e.g., square carriers for a robotic voice) of speech (Cohn 2013b). Other semi-productive classes may exist in different visual languages. JVL uses several affixes that are placed either on the forehead—like a sweat drop for anxiety in Figure 1a—or emerging from the nose—like bloody noses for lust or a bubble for sleepiness (Cohn and Ehly 2016). Might these constitute (or become) classes of "foreheadfixes" or "nosefixes"? Such analysis requires further research, but visual morphology clearly allows for both fully productive and semi-productive morphology, as in other modalities.

9 Empirical research

The growing literature on visual morphology over the past several decades raises several issues for future research, including cross-cultural variation, processing, and acquisition. I review this work below.

9.1 Cross-cultural variation

Visual morphology differs across cultures. This is particularly apparent in the conventions of one culture that appear opaque to those unfamiliar with that visual language. For example, without fluency in Japanese Visual Language, bloody noses for lust and bubbles out of a nose for sleep might seem perplexing. Just as the vocabularies of spoken and signed languages differ across the world, so too do those of visual languages. Research on such diversity can thus follow the lead of corpus linguistics research to investigate the variation, historical development, and typology of the lexicons of visual languages. This includes the variation between visual languages used in comics, those outside of comics, and their relations. For example, the emoji now popular in digital communication originally borrowed heavily from the vocabulary of Japanese Visual Language used in manga (Danesi 2016; Katsuno and Yano 2002). Some researchers subdivide between the visual morphology that appears specifically in the visual languages of comics (e.g., motion lines, pain stars), and those used in comics appropriated from outside those specific visual languages (e.g., hearts, dollar signs) (Forceville et al. 2014). Such categorization is often hard to distinguish, but increased corpus analyses could clarify these distinctions.

A growing literature of theoretical and corpus studies has already begun. Some work has attempted to simply characterize and categorize various lexical items both within and across comics (Cohn and Ehly 2016; Forceville 2011), especially word balloons and thought bubbles (Cohn 2013a; Forceville et al. 2010; Pratha et al. 2016). Similar work has focused on visual lexical items in the context of conveying conceptual metaphors (Forceville 2005; Shinohara and Matsunaka 2009; Abbott and Forceville 2011). Morphology has also been studied in visual languages outside of comics, such as in the sand drawings used by Central Australian Aboriginals (Wilkins 1997; Munn 1962, 1966, 1986/2016; Green 2014), and in older visual languages like those found on Mayan pottery (Wichmann and Nielsen 2016) and in the Old English Bayeux Tapestry (Díaz Vera 2013b, 2013a), among others.

Corpus research also has looked at morphology as a way to characterize the differences between visual languages. Cohn and Ehly (2016) found that most closed-class visual morphology used in 10 shonen (boy's) and 10 shojo (girl's) manga from Japan were similar, suggesting a shared visual vocabulary. However, different morphemes were used in varying proportions between genres, suggesting distinct sub-dialects. Such analyses can be extended across comics from other cultures and time periods (and contexts beyond comics) to explore to what degree different visual lexical items have spread across the globe and/or different time periods.

At the same time, cross-cultural studies can examine the degree to which there may be universal tendencies for systems to convey similar conceptual information. For example, does closed-class morphology in visual languages across the world tend to convey non-iconic information? While particular morphemes are no doubt culturally specific (even extended across cultures via globalization), certain meanings may be consistently conveyed across visual languages. Consider the sand drawings by Australian Aboriginals. Though they differ in most regards from the visual languages found in comics of the world, they do use a variety of conventionalized ways of drawing paths, somewhat akin to motion lines (Green 2014; Wilkins 1997/2016; Munn 1986/2016).

As corpus and annotation efforts grow, research can address questions of cross-cultural diversity and typology of the vocabularies of visual languages around the world and in historical contexts. It can also examine more carefully the development and history of changes in lexical items over time.

9.2 Processing

Research on the processing of various aspects of visual morphology goes back decades. Early work focused on the basic construal of meaning from these forms, especially with regard to the ages children begin to understand them (see below). Substantial work has been done looking at how kids understand carriers—like speech balloons and thought bubbles (Yannicopoulou 2004). This includes work suggesting that thought bubbles can be a successful intervention tool for teaching *theory of mind* to individuals with autism (Kerr and Durkin 2004; Parsons and Mitchell 1999; Wellman et al. 2002)—i.e., the idea that other individuals have thoughts that are different from one's own (Premack and Woodruff 1978).

Additional psychological research has looked at the understanding of motion lines. Some work has claimed that motion line understanding originates in basic aspects of vision, mimicking the "streaks" that are left behind in the visual system when viewing a moving object (Burr 2000; Burr and Ross 2002; Kawabe and Miura 2006, 2008; Kim and Francis 1998). However, recent research has shown that motion line understanding cannot be attributed to basic perceptual processing (Cohn and Maher 2015; Ito et al. 2010), and understanding this conventional representation is modulated by fluency in the visual language of comics (Cohn and Maher 2015). Measurements of electrophysiology have also suggested that the omission of motion lines evokes brain responses similar to those for incongruously reversed motion lines (Cohn and Maher 2015). Such findings suggest that, at least within the context of comics, motion lines do not just add meaning to an otherwise understandable representation, but rather they are an expected part of depicting events.

Research has also investigated more compositional aspects of visual morphology, such as the constraints on upfixes (Cohn et al. 2016; Ojha 2013; Newton 1985). This work has shown that upfixes need to be above a head, not beside it, and the upfix must "agree" with the facial expression (Cohn et al. 2016). For example, as mentioned above, a smiling face cannot be below a storm cloud. These constraints hold for both conventional and novel upfixes, suggesting that this is a semi-productive class of visual morphology, not simply memorized tokens.

Several questions have emerged as salient for future research: What are the cognitive mechanisms at work in compositional aspects of building meaning from monomorphs and visual affixes? To what extent do these processes overlap with the mechanisms operating in verbal and signed languages? How are these structures balanced with aspects of perceptual processing? And, to the degree that visual morphology differs between visual languages, how does experience with different visual morphology modulate their understanding?

9.3 Acquisition

An additional line of research involves how people learn this visual vocabulary. Most prior work on the acquisition of visual morphology has focused on the ages at which children are able to understand their meanings. In general, visual morphology appears to be understood better as individuals age (Nakazawa 2016), and may also be modulated by the frequency that those visual morphemes appear in comics (Newton 1985).

Nevertheless, the developmental trajectory of understanding visual morphemes may vary. For example, the meanings of speech balloons and thought bubbles appear to be understood by around 4 years old (Wellman et al. 1996). However, additional traits like loudness—denoted by jagged lines—appear to be understandable by even preliterate children (Yannicopoulou 2004). Motion line understanding progresses as children age, going from a fairly moderate understanding around age 6 to full understanding by age 12 (Friedman and Stevenson 1975; Gross et al. 1991; Carello et al. 1986; Mori 1995; Nakazawa 2016). This understanding also involves the shift from children interpreting them as physical elements (like wind) to fully symbolic conventions (Gross et al. 1991). This developmental trajectory differs from reduplication of body parts to show movement, which is understood even at earlier ages (Friedman and Stevenson 1975).

Thus, different visual morphemes do not appear to have a uniform timeline for the acquisition of their meanings. Future research can further examine these types of trajectories, but can also progress beyond interpretation alone to investigate compositionality and constraints on structure.

10 Conclusion

This chapter has examined the combinatorial aspects of visual lexical items, mostly between bound morphemes and isolatable forms (monomorphs). These combinations create meaning out of a balance between the intrinsic meaning of a representation with its spatial distribution using similar strategies as in spoken and signed languages: affixation, suppletion/substitution, and reduplication. This combinatorial meaning may involve a variety of semiotic reference types—possibly with inferential meaning drawing upon idiomatic or metaphorical knowledge. Such structure thus arises similarly between visual lexical items and those in the verbal and signed modalities. Drawing these parallels can hopefully inspire future work integrating research on visual languages into the linguistic sciences.

Acknowledgements

Kaitlin Pederson and Ryan Taylor are thanked for the brainstorming that contributed to this chapter, and Joost Schilperoord, Ray Jackendoff, and Charles Forceville are thanked for helpful feedback on earlier drafts.

References

Abbott, Michael, and Charles Forceville. 2011. Visual representation of emotion in manga: Loss of control is Loss of hands in *Azumanga Daioh* Volume 4. *Language and Literature* 20 (2):91-112. doi:10.1177/0963947011402182.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Audring, Jenny, and Geert Booij. 2016. Cooperation and coercion. *Linguistics* 54 (4):617-637. doi:10.1515/ling-2016-0012.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Burr, David C. 2000. Motion vision: Are "speed lines" used in human visual motion? *Current Biology* 10 (12):R440-R443. doi:10.1016/s0960-9822(00)00545-5.

Burr, David C., and John Ross. 2002. Direct evidence that "speedlines" influence motion mechanisms. *The Journal of Neuroscience* 22 (19):8661-8664.

Carello, Claudia, Lawrence D. Rosenblum, and A. Grosofsky. 1986. Static depiction of movement. *Perception* 15 (1):41-58.

Cohn, Neil. 2007. A visual lexicon. *Public Journal of Semiotics* 1 (1):53-84.

Cohn, Neil. 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, Neil. 2013a. Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica* 2013 (197):35-63. doi:10.1515/sem-2013-0079.

Cohn, Neil. 2013b. *The visual language of comics: Introduction to the structure and cognition of sequential images*. London, UK: Bloomsbury.

Cohn, Neil. 2016a. Linguistic Relativity and Conceptual Permeability in Visual Narratives: New Distinctions in the Relationship between Language(s) and Thought. In *The Visual Narrative Reader*, ed. Neil Cohn, 315-340. London: Bloomsbury.

Cohn, Neil. 2016b. A multimodal parallel architecture: A cognitive framework for multimodal interactions. *Cognition* 146:304-323. doi:10.1016/j.cognition.2015.10.007.

Cohn, Neil, and Sean Ehly. 2016. The vocabulary of manga: Visual morphology in dialects of Japanese Visual Language. *Journal of Pragmatics* 92:17-29. doi:10.1016/j.pragma.2015.11.008.

Cohn, Neil, and Stephen Maher. 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, Neil, Beena Murthy, and Tom Foulsham. 2016. Meaning above the head: Combinatorial constraints on the visual vocabulary of comics. *Journal of Cognitive Psychology* 28 (5):559–574. doi:10.1080/20445911.2016.1179314.

Culicover, Peter W., and Ray Jackendoff. 2012. Same-except: A domain-general cognitive relation and how language expresses it. *Language* 88 (2):305-340.

Danesi, Marcel. 2016. *The Semiotics of Emoji: The Rise of Visual Language in the Age of the Internet*. London: Bloomsbury.

Díaz Vera, Javier E. 2013a. Embodied emotions in Medieval English language and visual arts. In *Sensuous cognition explorations into human sentience: Imagination, (e)motion and perception*, eds. Rosario Caballero, and Javier E. Díaz Vera, 195-219. Berlin: de Gruyter.

Díaz Vera, Javier E. 2013b. Woven emotions: Visual representations of emotions in medieval English textiles. *Review of Cognitive Linguistics* 11 (2):269-284. doi:10.1075/rcl.11.2.04dia.

Engelhardt, Yuri. 2002. The Language of Graphics. Dissertation, University of Amsterdam, Amsterdam.

Fauconnier, Gilles, and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Fay, Nicolas, Mark Ellison, and Simon Garrod. 2014. Iconicity: From sign to system in human communication and language. *Pragmatics & Cognition* 22 (2):244-263. doi:doi:10.1075/pc.22.2.05fay.

Forceville, Charles. 2005. Visual representations of the idealized cognitive model of anger in the Asterix album *La Zizanie*. *Journal of Pragmatics* 37 (1):69-88.

Forceville, Charles. 2011. Pictorial runes in *Tintin and the Picaros*. *Journal of Pragmatics* 43 (3):875-890.

Forceville, Charles. 2016. Conceptual Metaphor Theory, Blending Theory, and other Cognitivist Perspectives on Comics In *The Visual Narrative Reader*, ed. Neil Cohn, 89-114. London: Bloomsbury.

Forceville, Charles, Elizabeth El Refaie, and Gert Meesters. 2014. Stylistics in comics. In *The Routledge Handbook of Stylistics*, ed. Michael Burke, 485-499. London: Routledge.

Forceville, Charles, Tony Veale, and Kurt Feyaerts. 2010. Balloonics: The Visuals of Balloons in Comics. In *The Rise and Reason of Comics and Graphic Literature: Critical Essays on the Form*, eds. Joyce Goggin, and Dan Hassler-Forest, 56-73. Jefferson: McFarland & Company, Inc.

Friedman, Sarah L., and Marguerite B. Stevenson. 1975. Developmental Changes in the Understanding of Implied Motion in Two-dimensional Pictures. *Child Development* 46:773-778.

Garrod, Simon, Nicolas Fay, John Lee, Jon Oberlander, and Tracy MacLeod. 2007. Foundations of representation: where might graphical symbol systems come from? *Cognitive Science* 31 (6):961-987.

Green, Jennifer. 2014. *Drawn from the ground: Sound, sign and inscription in Central Australian sand stories*. Cambridge, UK: Cambridge University Press.

Gross, Dana, Nelson Soken, Karl S. Rosengren, Anne D. Pick, Bradford H. Pillow, and Patricia Melendez. 1991. Children's Understanding of Action Lines and the Static Representation of Speed of Locomotion. *Child Development* 62:1124-1141.

Ito, Hiroyuki, Takeharu Seno, and Miyuki Yamanaka. 2010. Motion impressions enhanced by converging motion lines. *Perception* 39 (11):1555-1561.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray, and Jenny Audring. 2016. Morphological schemas. *The Mental Lexicon* 11 (3):467-493. doi:doi:10.1075/ml.11.3.06jac.

Katsuno, Hirofumi, and Christine R. Yano. 2002. Face to face: on-line subjectivity in contemporary Japan. *Asian Studies Review* 26 (2): 205 - 231.

Kawabe, Takahiro, and Kayo Miura. 2006. Representation of dynamic events triggered by motion lines and static human postures. *Experimental Brain Research* 175 (2):372-375. doi:10.1007/s00221-006-0673-6.

Kawabe, Takahiro, and Kayo Miura. 2008. New Motion Illusion Caused by Pictorial Motion Lines. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)* 55 (4):228-234. doi:10.1027/1618-3169.55.4.228.

Kerr, Sharyn, and Kevin Durkin. 2004. Understanding of Thought Bubbles as Mental Representations in Children with Autism: Implications for Theory of Mind. *Journal of Autism and Developmental Disorders* 34 (6):637-648. doi:10.1007/s10803-004-5285-z.

Kim, H., and G. Francis. 1998. A computational and perceptual account of motion lines. *Perception* 27 (7):785-797.

Koch, Walter A. 1971. *Varia Semiotica*. Hildensheim: Olms.

Lakoff, George. 1992. The Contemporary Theory of Metaphor. In *Metaphor and Thought*, ed. Andrew Ortony. Cambridge, 202-251. UK: Cambridge University Press.

Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by*. Chicago, IL: University of Chicago Press.

McCloud, Scott. 1993. *Understanding Comics: The Invisible Art*. New York, NY: Harper Collins.

Mori, Kazuhiko. 1995. The influence of action lines on pictorial movement perception in pre-school children. *Japanese Psychological Research* 27 (3):183-187.

Munn, Nancy D. 1962. Walbiri Graphic Signs: An Analysis. *American Anthropologist* 64 (5):972-984.

Munn, Nancy D. 1966. Visual Categories: An Approach to the Study of Representational Systems. *American Anthropologist* 68 (4):936-950.

Munn, Nancy D. 1986/2016. The Warlbiri Sand Story. In *The Visual Narrative Reader*, ed. Neil Cohn, 230-249. London: Bloomsbury.

Nakazawa, Jun. 2016. Manga literacy and manga comprehension in Japanese children. In *The Visual Narrative Reader*, ed. Neil Cohn, 157-184. London: Bloomsbury.

Newton, Douglas P. 1985. Children's Perception of Pictorial Metaphor. *Educational Psychology* 5 (2):179-185. doi:10.1080/0144341850050207.

Ojha, Amitash. 2013. An experimental study on visual metaphor. Dissertation, International Institute of Information Technology, Hyderabad, India.

Parsons, Sarah, and Peter Mitchell. 1999. What Children with Autism Understand about Thoughts and Thought Bubbles. *Autism* 3 (1):17-38. doi:10.1177/1362361399003001003.

Petersen, Robert S. 2011. *Comics, manga, and graphic novels: a history of graphic narratives*. Santa Barbara, CA: ABC-CLIO.

Phillips, Barbara J., and Edward F. McQuarrie. 2004. Beyond Visual Metaphor: A New Typology of Visual Rhetoric in Advertising. *Marketing Theory* 4 (1-2):113-136. doi:10.1177/1470593104044089.

Pratha, Nimish K., Natalie Avunjian, and Neil Cohn. 2016. Pow, punch, pika, and chu: The structure of sound effects in genres of American comics and Japanese manga. *Multimodal Communication* 5 (2):93-109.

Premack, David, and Guy Woodruff. 1978. Does the chimpanzee have a 'theory of mind'? *Behavioral and Brain Science* 4:515-526.

Sadock, Jerrold M. 1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago, IL: University of Chicago Press.

Schilperoord, Joost. 2013. Raising the Issue: A Mental-Space Approach to Iwo Jima-Inspired Editorial Cartoons. *Metaphor and Symbol* 28 (3):185-212. doi:10.1080/10926488.2013.768513.

Schilperoord, Joost. 2017. Ways with Pictures: Visual Incongruities and Metaphor In *Visual Metaphor; Structure and Process*, ed. Gerard Steen. Amsterdam: John Benjamins.

Schodt, Frederik L. 1983. *Manga! Manga! The World of Japanese Comics*. New York: Kodansha America Inc.

Shinohara, Kazuko, and Yoshihiro Matsunaka. 2009. Pictorial metaphors of emotion in Japanese comics. In *Multimodal Metaphor*, eds. Charles Forceville, and Eduardo Urios-Aparisi, 265-293. Berlin/ New York: Mouton De Gruyter.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.

Walker, Mort. 1980. *The Lexicon of Comicana*. Port Chester, NY: Comicana, Inc.

Wellman, Henry M., Simon Baron-Cohen, Robert Caswell, Juan Carlos Gomez, John Swettenham, Eleanor Toye, and Kristin Lagattuta. 2002. Thought-bubbles help children with autism acquire an alternative to a theory of mind. *Autism* 6 (4):343-363.

Wellman, Henry M., Michelle Hollander, and Carolyn A. Schult. 1996. Young children's understanding of thought bubbles and thoughts. *Child Development* 67:768-788.

Wichmann, Søren, and Jesper Nielsen. 2016. Sequential Text-Image Pairing among the Classic Maya. In *The Visual Narrative Reader*, ed. Neil Cohn, 282-313. London, UK: Bloomsbury.

Wildfeuer, Janina, and John Bateman. 2016. Linguistically Oriented Comics Research in Germany. In *The Visual Narrative Reader*, ed. Neil Cohn, 19-66. London, UK: Bloomsbury.

Wilkins, David P. 1997. Alternative Representations of Space: Arrernte Narratives in Sand. In *Proceedings of the CLS Opening Academic Year '97 '98*, eds. M. Biemans, and J. van de Weijer, 133-164. Nijmegen: Nijmegen/Tilburg Center for Language Studies.

Wilkins, David P. 1997/2016. Alternative Representations of Space: Arrernte Narratives in Sand. In *The Visual Narrative Reader*, ed. Neil Cohn, 252-281. London: Bloomsbury.

Yannicopoulou, Angela. 2004. Visual Aspects of Written Texts: Preschoolers View Comics. *L1-Educational Studies in Language and Literature* 4:169-181.