**Nouns for Visual Objects: A hypothesis of the vision-language interface**

Paolo Acquaviva and Francesco-Alessio Ursini

**1 Introduction**

Does the English word *apple* correspond to a concept APPLE? Disregarding the uncertainty introduced by “correspond to”, a positive answer seems natural, yet problematic. It is natural because the word identifies a type of object, which speakers are disposed to call apples, and “concept” seems an adequate description for the total knowledge associated with this type of object. However, concepts cannot be just the content of words. The store of things, events, relations, facts, that we can think of, and that we perceive as discrete aspects of experience, are not bounded by our vocabulary. Otherwise, there would be no concepts without language, for example in infants. Reducing concepts to word content is made even more problematic by linguistic diversity. The English *apple* and the French *pomme* enter into different networks of associations. For instance, only the French noun is also employed in describing the potato, yet it is far from obvious that they should correspond to different concepts for that.

The relation between concepts and what is often known as lexical concepts raises many questions. This paper addresses them in relation to the specific sub-domain of physical objects identified visually, or “visual objects” as we will call them. The visual-cognitive basis makes it possible to anchor many of their fundamental properties outside language, escaping the typical risk of circularity that threatens such investigations. Because a suitable theory can provide a language-independent characterization of a visual object, we can relate it to a theory of what a noun can mean in a natural language and examine if the two are isomorphic.

Our contribution necessarily takes the form of a hypothesis. This because there is no “objective”, definitive account of how we conceptualize visual percepts, nor a generally accepted theory of what concepts a noun can and cannot denote in natural language. We will therefore propose in sections 2 and 3 two answers to these questions, rooted in specific approaches to language and cognition. These approaches can be defined as computational and information-theoretical, offering a mind-internal view on the problem at hand. They rest on bases which have been justified independently, a fact we document by engaging extensively with the relevant literature. This does not mean that they are the only correct accounts of how vision and language construe entities. However, it does show that they are not an arbitrary choice, in function of the intended conclusion.

The summaries we offer of how the mind encapsulates information about visual objects and nouns are contributions in their own right, but our main claim is about the way the mind connects them. This is is the goal of the formalization offered in sections 4 and 5. As a result of what (we claim) can be conceptualized as an object on visual and on linguistic bases, we will be in a position to constrain the possible range of visual objects expressible as nouns, ruling out a priori thinkable candidates and so making an empirically falsifiable hypothesis about the boundaries of linguistically encapsulated concepts.

**2 The ingredients of visual objects: definitions, types and processes**

We start by presenting a hypothesis of visual objects (Section 2.1) and showing how they capture our understanding of visual information (Section 2.2).

Our hypothesis takes the form of a computational, internalist and representational view on vision. We choose a computational, internalist view because works in this line of research offer more detailed accounts of visual features, and address questions akin to our main questions (cf. Pylyshyn 1984: Ch. 1–2; Scholl 2001; Collins 2017). Alternative views with an ecological, externalist bent are certainly possible. We discuss them when we discuss the notion of “affordance” (Gibson 1966; 1979).

*2.1 Objects and types*

The human visual system processes the continuous stream of visual information about the real world by unconsciously segmenting it into discrete spatio-temporal units (Michotte 1963, Ullmann 1979; Marr 1982: 10-34). These units are then identified as “objects” insofar as they stand for mental categories that guide our conscious understanding of the world. Research in object recognition has accumulated growing evidence that specialized visual sub-systems regulate the recognition of certain properties of visual information. However, the underlying principles governing recognition processes seem to be the same across these sub-systems (Scholl 2009). This suggests that visual objects can be identified as the results of the mental computations and processes that connect different types of visual information. In other words, a blue ball can be identified *qua* object of an identifiable colour and shape.

A first model that attempted to capture this computational principle is *Multiple Object Tracking* (Pylyshyn 1984: 34–40, 1989). This model assumes that objects correspond to the mapping of *visual tokens* to the *visual types* they instantiate. Visual types can be combined to form abstract representations of individuals’ experiences classified via types. The result of this mapping is a visual index representing a mind-external object, known as a *Finger of INSTantiation* or FINST (Pylyshyn 1989: 69; Scholl, Pylyshyn and Feldman 2001). A FINST thus acts as a complex “mental finger” that individuates an object via the properties/types it instantiates(or realizes). In the case of the blue ball, a visual stimulus/token instantiating the blue colour type and the round shape type is FINSTed, or indexed, in virtue of instantiating these two visual types over time. This internal representation of an object is in turn classified as one type of representation for a *BALL* concept.

Much research on object recognition has focused on the sub-types involved in these computations. For instance, it was suggested that *motion* is a visual type that allows us to identify objects (Michotte 1963; Ullmann 1979; Tremoulet and Feldman 2000, 2006). Motion is often associated to agentive behaviour and to the ability to move independently, whether this agency is apparent or real. If one observes several dots moving after a single dot on a screen, one can perceive the former as a group of dots chasing the lone dot (Gao, Newman and Scholl 2010; Scholl and Gao 2013; Scholl and Tremoulet 2000).

 The physical properties of objects, collectively known as *textons* (Julesz 1981), act as a sub-domain of visual types that guide object recognition. Texture (Zhu, Guo, Wang and Xu, 2005), colour (Alvarez-Fernandez and Varnell 2011), material constituency (Van Marle and Scholl 2003), and shape (Ben-Shahar, Ohad and Zucker 2007) also are proposed types. *Shape* types have been investigated and are a specific sub-domain of object recognition. Initial representational models (e.g. Marr 1982’s three-layered model) have been developed into more complex ones. Some, such as *Recognition by Components,* involve multi-layered representations identifying and relating *geons*, complex shape types (Biederman 1987; Hummel and Biederman 1992, Riesenhuber and Poggio 1999, 2000).

Objects can be individuated as quantities or *units* of a possibly underspecified type (Pylyshyn 1984, 1989). Thus, several dots moving in the same direction can also be individuated as a single complex unit or group (Scholl, Feldman and Pylyshyn 2001; Pylyshyn 2004, 2006). If a group of objects splits into its constituent units, each unit can be identified as a distinct object (Mitroff, Scholl and Wynn 2004). Obejcts can be tracked according to the quantities they instantiate, e.g. when observers receive specific instructions to do so (“nine” or “most” balls on screen: Pietroski, Lidz, Hunter and Halberda 2009).

Objects can also be individuated via their potential functions for a subject in an environment, or *affordances* (Gibson 1966, 1979: Ch. 1). Thus, the precise function of an object becomes context- and subject-sensitive. A man can “see” a hammer as a tool for hitting nails; a child, as a toy to break stones with. Recent works on the notion of affordance suggest that subjects may develop stable, mind-internal representations of affordances over time (Anderson *et al.* 2002, Wells 2002, Whitagen *et al.* 2012). Both functions can then be memorized and internalised as features assigned to the object “hammer”.[[1]](#footnote-2) Other visual types can also have an externalist, ecological interpretation. One can interpret a “colour” type as a feature of an object in the environmnent that an agent experiences. The *conundrum* that arises in such a case is how an agent would be able evaluate whether such an object is novel or old, and thus requires a novel name for it or not.

Overall, object types can encompass various conceptual domains that organize visual information into higher-order, mind-internal structures. Even if several domain types exist that observers can use to classify visual tokens, the process of individuation seems streamlined. When a visual token is recognized as instantiating at least one visual type, this produces a visual object as a mental representation for this successful representation. The visual computations that permit observers to recognize objects also follow formal constraints establishing possible relations between objects and their underlying types. We discuss four such constraints (two spatial, two temporal), to illustrate their underlying mechanisms.

Two spatial constraints for object individuation are *connectedness* and *mereology*. In the context of our system, these constraints are spatial~~,~~ because they determine how objects are individuated in the visual array via their features. Their geometrical relations are consequently inferred. Connectedness is a property that establishes that objects and their parts occupy connected, possibly contiguous regions of the visual space. Thus, a red spot on a blue car may signal a distinct part of this car, since its colour is not the same of contiguous regions of the car (Wolfe and Bennett 1997). Mereology is a property that establishes that objects can also have parts also individuated as distinct objects. For instance, an ice-cream can be recognized as a more complex object that includes at least one scoop and a cone as parts (Marr 1982: 222-230; Hummel and Bidermann 1992; Riesenhuber and Poggio 2000).

Two temporal constraints are *persistence* and *cohesion*. These constraints are temporal in nature because they determine how the features individuating objects are activated over the temporally extended process of object recognition. Persistence states that the types that identify an object remain stable over time unless some change is detected (Scholl 2009: 680-682). Cohesion states that types can combine, as far as they do not conflict in their nature (e.g. Scholl 2007, 2009: 684). A representation of an ice-cream involves a cone and scoop over time, unless some change (like the scoop melting) is detected.

We define spatio-temporally stable object representations as *object files,* mental representations that can also be updated, as new information is processed (Kahneman et al. 1992; Zacks 2004; Zacks et al. 2007). Observers can maintain and update object files by instantiating higher-order representations for classes of objects (Pylsyhyn 2004; Scholl 2007, 2009). Object files as internal representations of mind-external entities are elements computed in short-term memory. Object *concepts* as representations of object classes represent information stored in long-term memory (Biederman 1987; Cowan 1995, 2005).

Object concepts, as long-term memory items, can involve complex representations for abstract classes. For instance, the “skateboard” object concept can correspond to a representation for any object having certain key parts (four wheels, a table) and functions (use as a locomotion tool). Observers can use this tacit knowledge to complete and update these representations over time. Thus, an observer witnessing a skateboard missing a wheel by accident will still be able to consider this object as a skateboard (see Scholl and Pylyshyn 1999; Serre, Wolf and Poggio 2005; Feldman 2007; Scholl 2001, 2009).

The upshot of our discussion is as follows. Objects, or more accurately object files, are intermediate representations emerging from the instantiation relations holding between more basic elements. These are visual tokens and visual types, sensory impressions and the mental representations they activate. Object files that become spatio-temporally stable trigger a higher order instantiation relation between the object and the object concept they instantiate. In sum, object files represent temporary classification of the entities we individuate as we make sense of the visual stream. When these representations become stable, we connect them to abstract concepts we have memorized to classify and understand the world.

*2.2 A Theory of visual objects*

We now make formally precise the notions that we discussed in the previous section. We take Multiple Object Tracking theory as a guiding model (Pylyshyn 1984, 1989), although our notation will diverge from this framework when needed. In this theory, an object is conceived as an index (Finger of INSTantiation or FINST) associated to the relation between a visual token and the visual type(s) it instantiates. Formally, a FINST is defined as *FINST(token,type)=object* (Pylyshyn 1989:75–76).This definition says that an object is individuated when a visual token instantiates a visual type. The crucial extension we make to this definition is that FINST and “object file” are near-synonymous labels for the same formal notion, given their similar function (but see Pylyshyn 2009 for discussion).

The next step is to define how we represent object concepts. This aspect is usually underdeveloped in the theories discussed so far, but can be made precise by implementing a type theory in our analysis. Before we proceed, a theoretical *caveat* is due. The formal representation for the relations we discuss in this section (e.g. the *part-of* relation) may not capture their use and interpretation in the cited works (cf. Marr 1982: 228–240). However, this is the case because definitions in most previous works take an informal, and often theoretically ambiguous approach that we avoid via the use of this formalism. Within formal approaches to the semantics of natural languages, several such approaches exist, like situation semantics (Ginzburg 2005), Type Theory with records (Cooper 2004), Type Logical Composition (Asher 2011). A type theory in which basic types can be combined into more complex types (e.g. Asher 2011), will be germane to our explanatory goals.

We first assume that object concepts can be defined in an extensional and in an intensional manner. A concept *C* can be extensionally conceived as the set of object files instantiating certain properties. Thus, the object concept of red things is the collection of object files *o*, *o’*,*o’’*…whose individuating type is *red*. We represent object files via lower-case letters distinguished by superscripts. We then represent concepts based on types by means of upper-case letters spelling out the concept’s constitutive type. Thus, the concept “red” is represented as *RED*, and *{o, o’, o’’}* = *RED* is the relation stating that a concept is defined as the set of objets files with a given type (here, *red*). A concept can also be intensionally conceived as a *membership* relation between each object file and the concept so defined (so that we have *o ∈ RED*, with “*∈* “ representing the membership relation).

Our second assumption, adopted from various frameworks (e.g. Asher 2011), is that types and consequently concepts can be combined into complex types and complex concepts. Since object files can instantiate several types at once (for instance we can have white, oval balls as “tools” used in rugby), they can also be classified via concepts that represent and store such complex combinations. For type combinations, we assume that types and type relations can be modelled via a *Boolean algebra*, a structure in which the three operations *meet*, *join* and *negation*, and the relation *part-of* are defined.

Our choice of these algebraic structures is based on the fact that several works we reviewed so far also follow this view. For instance, in Recognition by Components theory complex shapes are modelled as objects instantiating joins of simpler forms (cf. Hummel and Biederman 1992). FINSTs in Multiple Object Tracking Theory are modelled as conjunctions (i.e. meets) of visual types (Pylyshyn 2004). The part-of relation is often used model relations between objects, as well as the spatial regions they occupy (e.g. Marr 1982: 225–30). Thus, the use of a formal algebra of types is also consistent with previous literature.

We now make these notions precise. To make the discussion maximally transparent, we represent this and the other operations/relations by representing how types can be applied to tokens. The operation *meet*, represented via the symbol“⊓”, is defined as an operation that takes types and returns their common type(s), defined as “*a(x)⊓b(x)*”. Thus, *blue(x)⊓ball(x)* is the meet of the visual types that represent a blue ball, via a visual token *x*.[[2]](#footnote-3)

The operation *join*, represented via “*⊔*”, takes types and returns their unified type(s), defined as “*a(x)⊔b(x)*”. Thus, an ice-cream shape corresponds to the join of two more basic shapes, that of a cone and a sphere: we have *cone(x)⊔sphere(x)=ice-cream(x)*. The operation *negation* “*¬*” is defined as an operation that takes a type as an input and returns its complementary type as an output. If square is divided into two distinct parts, one black and the other white, then the colour type of one part will be the complementary type of the other part. We thus have *black(x)=¬white(x)*. The part-of relation “*⊑*” (i.e. “*a(x)⊑b(x)*”) is defined when the two identities *“a(x)⊓b(x)=a(x)*”and “*a(x)⊔b(x)=b(x)*” hold. That is, type *a* is part of type *b* if their meet is the (sub-)type *a*, and their sum is the (super-)type *b.* A second relation is the identity relation. Two object files are identical (i.e. we have *a(x)=b(x)*) if their meet and join are also identical (i.e. we have *a(x)⊓b(x)=a(x)⊔b(x)*).

Let us now turn to object files. According to these definitions, an ice-cream can be represented as the join of two shapes: we have *cone(x)⊔sphere(x)=ice-cream(x)*. An ice-cream object file, call it *b*, is defined as *b=(cone(x)⊔sphere(x))*. An index *b* identified with an instantiation relation between a visual stimulus *x* and a join of two shape types, *cone* and *sphere*. This simplified analysis already captures the fact that object files are established when a visual token *x* is recognized as instantiating specific complex types. A similar analysis can be extended to meet types. For instance, a blue ball can be defined as the object file *d=(blue(y)⊓sphere(y))*. The object file *d* is identified with a visual token *y*, instantiating the visual types *blue* and *ball* at the same time, hence the use of the meet operation. By contrast, the object file corresponding to an ice-cream cone is modelled by the union of two types, corresponding to its distinct visual components.

For the role of negation and the part-of relation, we need to discuss more examples. Since we now have a definition of object file, we can define the notion of object concept. Consider a case in which an agent sees a blue and a red ball. Both are “ball” objects, even if they have distinct colours. Their representations can be formalized as *d=(blue(y)⊓sphere(y))* and *f=(red(z)⊓sphere(z)),* respectively. Both are also balls, whether their colour is blue, red or some other colour that an agent has memorized for balls. In other words, a “ball” object concept is a mental representation in which objects are classified together with respect to one type (here, a shape type), but are left *underspecified* with respect to other types (here, colour). Underspecification holds when the value assigned to a type can be ambiguous in context (cf. Pylyshyn 2004; cf. Asher 2011: Ch. 3; for a linguistic definition). In the case of object files, a colour feature may be temporarily underspecified. An observer may first recognise an object as a ball, and then determine its specific colour (e.g. blue, say in bad light).

Let us move to concepts. For concepts, we can use a slightly different and more compact notation. For instance, The object files *d=blue(y)⊓sphere(y)* and *f=red(z)⊓sphere(z)* may be the minimal members of the concept *BALL*, which can be defined as a complex concept representing a spherical object of any colour. Thus, the concept of *BALL* can be represented as the join *BALL=(SPHERE⊔COLOUR)*, with *SPHERE*,the concept derived from type of spheres, and *COLOUR* being the complex (join) concept of all *colour* types (e.g. we have *col=red⊔green⊔blue…*, hence *COLOUR=…*). It is also possible to use a variant notion used in the literature on Boolean operators to represent concepts (e.g. Landman 1991: Ch.2–4; cf.also Chierchia 1998 on kinds). We thus have *CONCEPT=(⊔file)*. This reads: an object concept is the join of all files representing a given object type (e.g. *ball(x)⊔ball(y)⊔…ball(z)*).

A first consequence of this definition is that an object concept *cannot* be instantiated via a single object file, for such an object would then involve the actual, physical instantiation of several distinct and possibly conflicting types at once. If we take *DOG* as a visual concept with specific identifying types, it is still possible to distinguish between *beagle* and *dachshund* sub-types/concepts, since aspects of their bodies differ considerably (see Riesenhuber and Poggio 1999). Therefore, our system proves that a concept such as *DOG* represents the combination (union) of more specific objects or “variants” that qualify as dogs because of their similar, and yet distinct visual properties (types, in our parlance).

A second consequence is that object files play two roles in the system of object recognition. From a bottom-up perspective, they represent information about external entities (e.g. their shape and colour). From a top-down perspective, they represent a specific instance or case of a concept. This entails that types representing visual features can be conceived as “primitive” categories. Under this view, learning and memorizing a concept is an ongoing process of “adding” novel information of what count as e.g. a ball (cf. Margolis and Laurence 1999; Murphy 2002: Ch. 2), and how different but related types can co-exist under this representation (Murphy and Ross 2010; Shuwairi, Bainbridge and Murphy 2013). Concepts, including object concepts, can develop over time.

We can show how our formal system works by showing how it can prove the recognition of an “incomplete” object, e.g. the skateboard mentioned in the previos section. If a *skateboard* object concept corresponds to the combination of a table and four wheels as “prototypical” shape, then a skateboard object file lacking a wheel (e.g. the front, left one, call it *w1*) will correspond to an object file instantiating all but one of the types making up this shape. The relation between the type that this partial skateboard and a skateboard object concept can be defined, in simplified manner, as in (1), which includes a partial skateboard object file (we drop capital letters and use types to identify objects, for readability):

(1) *partial\_skateboard(x)=*

 *=((table(x)⊓w2(x)⊓w3(x)⊓w4(x))⊑(table(z)⊓w1(z)⊓w2(z)⊓w3(z)⊓w4(z))*

 *⊨partial\_skateboard(x)∈SKATEBOARD*

The “partial skateboard” file involves the instantiation of most, but not all the types belonging to the “whole skateboard” concept. Object recognition is successful because of this relation holding between types and the object files they define. We represent this fact via “*⊨*”, the instantiation relation, a notion that will fully be defined in Section 4.

Overall, the notion of objecthood and of visual object files and concepts suggests that vision is a mediated process, in which the classification of visual stimuli plays a key role. The mental ontology that underpins visual objects reflects mind-internal principles regarding how visual information is processed and organized. Object concepts act as mental objects that guide this classification process even if the visual input is partial, or it involves apparently incomplete or underspecified information. The question that now arises is how this visual ontology is related to a linguistic one, specifically to an ontology of nouns.

**3 Nouns and nominals: the linguistic conceptualization of objects**

We now turn to the linguistic encapsulation of object concepts. We first outline the motivation for a theory of nouns as grammatical objects (3.1), then we focus on the relation between nominality and the denotation of entities (3.2), and outline a particular (and unconventional) way to understand this relation as a naming relation between a noun and an object concept (3.3). Then, section 3.4 makes precise the grammatical structure we posit for nouns and for nominals, while 3.5 clarifies the status of grammatical basic elements in the construction of conceptualization. Sections 3.6 and 3.7 make explicit the relation of nominals with the conceptual domain and detail the constraints this system imposes on the possible nominal encapsulation of entity concepts.[[3]](#footnote-4)

*3.1 Constraining noun content*

It may seem natural to view nouns as semantic atoms, where lexical semantics helps understand their content, but does not aim to explain what it can be. In fact, this conception derives from the relative underdevelopment of lexical semantic research on nouns, as opposed to verbs or deverbal nominalizations. Decades of work have shown that verbal meanings correspond to specific profiles in terms of Aktionsart, aspect, argument structure, further specified in terms of causation, agentivity, or affectedness (Binnick 2016; Krifka 1998; Rothstein 2004,;Borer 2005b; Ramchand 2008; among many). It is only recently that the fundamental parameters of noun semantics have come into focus.

 We adopt a constructionist perspective (Borer 2005a; Fábregas and Scalise 2012), following the insight that the content of nouns is shaped by knowledge of grammar. Noun meaning, like all lexical meaning, is modelled via syntactically complex grammatical objects, rather than structured templates, which interpret syntactically atomic lexical items. Nouns are thus pieces of noun phrase structure. Hence, the investigation of nouns shades into the tradition of studies into noun phrase structure (see Rijkhoff 2002; Borer 2005a; Svenonius 2008; Wiltschko 2012; Déchaine et al. 2014). It is not our aim here to argue for this choice, which shares with lexicalist approaches the goal of modelling context-invariant noun content in a way that predicts possible and impossible readings. We simply hold that a constructionist approach offers an explicit, empirically justified model of noun structure. The lack of a theoretical distinction between lexical nouns and grammatically constructed structures avoids the pitfall of severing morphological words and larger expressions. This lack of a categorical distinction from a grammatical viewpoint is desirable, given the factual observation that concepts are expressed both by syntactically simplex and complex structures.

*3.2 Nouns and object types*

If verbs refer to actions, processes, states, and generally eventualities, nouns refer to “things”, including concrete and visual objects (see Baker 2003 and Croft 2000 for different approaches sharing this starting point). This apparently innocuous statement conceals an ambiguity which the constructionist approach brings into the open. What really refers to actual, concrete objects, are not nouns per se, but nouns as heads of noun phrases in a referential use. Predicatively used nouns as in (2), nouns denoting roles or guises in constructions as in (3), but also nouns inside compounds like those underlined in (4), illustrate the fact that nominals (bare nouns and phrases) can have a non-referential value:

(2) a Mary is [a professional clown].

 b Mary is [a valued friend and colleague].

 c I consider you [my best friend].

 d He returned home [a happy father].

(3) a Henry was appointed [chairman].

 b Henry went to the party disguised as [a linguist].

(4) kill-joy, turn-coat , wind-fall, tape-worm

The bracketed constituents in (2)-(3) predicate properties of their subjects. However, it is the subjects that identify a referent, unlike identificational sentences like *Mary is the clown*, where two referents are stated to be the same. In (4), the underlined part does not refer to any one joy, coat, wind, or worm in particular, because it is the whole compound that identifies a particular referent. We will adopt the view that it is full noun phrases, and not simple nouns, which can refer properly. In a view of reference as a relation that speakers establish with some entities, it is through whole noun phrases that speakers identify specific “particulars”, or irreducible individuals (Strawson 1959). This includes phrases that correspond to a single noun, without modifiers or determiners.

In *she saw flashes*, *[flashes]* is a one-word noun phrase, because it paradigmatically alternates with expressions like *[some flashes]* or *[those two flashes]*. This phrase expresses the theme argument and refers to the object seen by the subject. The noun *flashes* by itself denotes the type of object referred to by the phrase; in the terms of section 2, noun phrases express object files, and nouns express object concepts. This clear theoretical choice provides the essential link with the preceding discussion of visual objects. There, we saw how visual types and tokens are interpreted as objects by being mapped onto object files. However, a noun can only be linked to a concept. It is only at this level that noun semantics and the content of visual sensory information are related, not at the level of visual properties or of object files. Grammar does not replicate the components of (visual) perception.

*3.3 Nouns, properties, and kinds*

The notion of kind current in formal model-theoretic semantics is obviously very closely related to the notion of object concept defined in section 2. Its main function since it was introduced by Carlson (1980) has been that of modelling the abstract denotation of terms like *gold* or *bears* in generic sentences like *gold / bears abound in this area*. A kind in this sense is an abstract entity instantiated by particular objects, as in the case of object concepts arising from object representations, but it is understood as the intensional abstraction of a property. This notion can be defined as follows.

In the influential formalization of Chierchia (1998), given a property *P*, the kind *K* is a function which for each possible world returns the maximal sum of objects true of *P* in that world. A kind is generally construed as a generalization from individual objects. A nominal kind is the union of all nominals representing a given noun type (e.g. *nominal(1)⊔nominal(2)⊔…⊔nominal(n)*). If a concept can be represented as the union of all relevant files (i.e. *⊔(file(n))=CONCEPT*), then a kind can the represented as the union of all nominals and their denotations (i.e. *nominal(n)=KIND*). [[4]](#footnote-5)

This makes it natural to view kinds as secondary formal constructs definable from any class of entities sharing a property. One example is the class of individuals with the same height, in the interpretation of degrees as kinds (Anderson and Morzycki 2015). Kinds in this sense are then secondary entities automatically definable based on any shared property. We must clarify that we restrict our attention to the kinds that reflect conceptual objecthood. Other types of kinds, while potentially relevant for semantic analysis, are not relevant.

Most work in formal semantics assumes nouns to denote properties of entities, so that *dog* would be interpreted as a predicate ***dogʹ*** true of individual dogs; its extension is a set, just like the extension of adjectives like *big*. We follow an alternative view, argued for by Krifka (1995), Zamparelli (2000), and Mueller-Reichau (2006), which holds that nouns denote primarily kinds, and only secondarily, through the addition of grammatical determinants, are they turned into predicates true of instances of kinds. We assume that the innermost part of the noun phrase denotes a kind as an entity of logical type *e* (individuals, cf. Lasersohn 2018), not a *<e,t>* property (a predicate) true of kinds (contrast Cyrino and Espinal 2015).

When we discuss types in the sense of “logical types”, we refer to a semantic literature in which ontological matters are to an extent simple. The type *e* represents any entity, and *t* is a type that represents a truth-value, e.g. whether it is true that an entity belongs to a certain category or not. Works like Asher (2011) propose that the type *e* can have a rich inventory of sub-types. Entities of type *e*  can actually belong to sub-types such as *object*, and be further specified for (sub-)sub-types such as *shape*, *colour* and so on (cf. also section 2). For our purposes, the alternative conception of kind that we propose here has a key function. It shows how the nominal system can “prove” the existence of abstract notions from noun tokens and types. We thus do not need to discuss nominal type in the thorough manner of Section 2.

A kind understood in this sense is an underived individual, posited because the workings of natural language (or rather our hypotheses about them) make it necessary, much along the lines of Carlson’s (1980) original conception. They are not merely the result of abstracting intensionally over a previously given domain of object-level entities. Carlson (1980: 60) noted that kind-referring expressions align with proper names and contrast with quantified noun phrases in that they can be arguments to the predicate *to be so-called*:

(5) a Slim is so-called because of his slender build.

 b The cardinal is so-called because of its colour.

 c Cardinals are so-called because of their colour.

 d \* Some / All / Most cardinals are so-called because of their colour.

However, common nouns too can act as naming labels, in a naming construction like (6) (see Acquaviva 2018). Here, what is said to bear the name ‘water’ is not the concrete portion of matter ostensively identified by the indexical *this*, but the kind it instantiates:

(6) This is called water.

We therefore assume that the minimal syntactic object that qualifies as a noun denotes a kind and names an object concept, and does so rigidly. Rigidity of reference entails that a noun names the same entity across all possible worlds (Kripke 1972). This also entails that the object concept so named is defined across possible worlds, and thus cannot be an entity limited to some worlds only or identified by indexicals. Therefore, our hypothesis predicts that no lexical noun can define an object concept like “the interval from a to b on this line”, even though this is certainly thinkable.

*3.4 Nouns as noun phrase substructures*

There are several ways to construe the notion of “minimal syntactic object that qualifies as a noun”, depending on the framework adopted. Our interpretation foregrounds the role of grammar in shaping what speakers know about lexical items. It thus dovetails with frameworks like Distributed Morphology (Embick and Marantz 2008) and Borer’s (2005a, b) “exoskeletal” model, where the same grammatical mechanisms govern the construction of lexical items and of syntactic phrases.

 We thus follow Borer (2005a) and the consensus view in Distributed Morphology (see Harley 2014). Hence, we interpret a “minimal syntactic object that qualifies as a noun” as a syntactic *root*, a category-free syntactically unanalyzable element, combined with a grammatical determination defining a nominal. In Distributed Morphological terms, this determination corresponds to a discrete [n] head. We leave open the possibility that the nominalizing information may consist of syntactically visible features, rather than a separate head (see Adger 2013 for a model of syntactic projection). We notate this nominalized root [N root]. In general, when no other element contributes to identifying a kind, this minimal noun denotes a kind, whose counterpart is an object concept, a point we fully explore in Section 4.

 Above [N root], further grammatical determinations turn the entity-denoting kind-name into a predicate, true of instances of the kind. This generalizes to all languages the kind-to-predicate operation that Chierchia (1998) posits for nouns in classifier languages like Chinese, when they combine with determiners. In this intermediate stage nouns acquire the grammatical properties which determine their behaviour as count or mass. Number has its encoding here. This holds not only as a determination about numerosity in terms of one-many (or also ‘a few’, or ‘uncountably many’: Harbour 2014), but also as the characterization of the denotation domain as simplex or having a complex internal structure. In line with much recent work, we do not posit a simple two-way switch between countable domains with discrete atoms and mass atomless domains,. Rather, we pose a set of determinations that distinguishes nouns whose denotation is dimensionally specifiable from those whose denotation is not, although they still denote spatially extended entities (Zhang 2012, 2013).

 Since our goal is not that of developing a detailed theory of noun and Determiner Phrase structure (to call the extended noun phrase by the label that is usual in syntactic studies), but of relating a specific approach to a hypothesis of vision, we will avoid more discussion and schematize the structure we assume in (7):

(7) QP/DP quantifiers, demonstratives, articles

 Quant/Det NumberP numerals, some plurals

 Number UnitP count classifiers, some plurals

 Unit SizeP shape classifiers, dimensional modifiers

 Size [N root] noun classifiers

 N root

 The highest layer in this structure hosts information about categories like definiteness, specificity, and all that connects the DP referent to its context. As mentioned before, we take this to be the locus of reference, via the determination of specific object files. The projections in the middle modulate the various aspects that constitute countability, distinguishing between a part structure arranged into discrete fixed-size units, which permits counting, and a division into discrete but not directly countable elements. Grammatical number and most classifiers fall into this layer (Svenonius 2008). The innermost projection, notated [N root], defines the nominality of a root, which may be morphologically signalled by specific morphemes (gender being a natural candidate) or also by those classifiers which accompany nouns independently of the presence of counting, measuring, or quantifying structures.

*3.5 Grammatical structure and conceptual structure*

It seems natural to interpret the structure of DP as a reflection of conceptual structure. However, the consensus is limited to associating the peripheral, intermediate, and innermost regions respectively to quantifying and relating the referent in context, to characterizing its denotation domain (including quantity and numerosity), and to classifying it. Rijkhoff (2002:50–53) established a direct connection between grammar and non-linguistic cognition and perception, by interpreting the features underlying his typology of countability in terms of “shape” (whether the referent has a definite outline) and “homogeneity” (whether the property true of the referent also applies to larger entities that include it). This does not amount to a reduction of grammatical categories to perceptual ones, since a “property designated by a noun” must be “coded” as having or lacking the two features (Rijkhoff 2002: 50, 53). However, the features are defined informally, using natural language as metalanguage in concrete perceptual terms. This approach makes them problematic for all abstract nouns like *idea* and for nouns like *thing*, which impose a count reading without any information about shape or internal structure.

 Most analyses envisage a much more indirect relation between the categories of experience and the content of grammatical features. In the case of countability, “homogeneity” has been interpreted via formal tools (Link 1983; Chierchia 1998, Krifka 2004; Borer 2005a; Rothstein 2010). They have brought out the importance of nouns like *segment* (which is count, but whose subpart is still a segment) or *furniture* for a theory that aims to cover all nouns. Researchers have also documented the need to distinguish the unitization needed for counting from the reference to discrete bounded objects (see the contributions in Massam 2012). This matches the distinction between numeral/counting classifiers and those classifiers that package a domain into discrete elements typically by qualifying an object’s appearance (Svenonius’ 2008 unit and sort classifiers; see recently Fassi-Fehri 2018 for this distinction from a non-classifier language).

Importantly, the properties made salient by classifiers are not necessarily immanent in the lexical semantics of a given noun. Therefore, they cannot be dictated by properties of its denotation. In the Australian language Yidiny ‘a piece of hot charcoal’ (*nirgil*), can be described as *buri* (cl:fire) or as *wirra* (cl:movable object), and a plot of ground either as *jabu* (cl:ground) or as *bulmba* (cl:habitable)’ (Dixon 1977: 203, cited in Aikhenvald 2003: 84). The change in perspective brought about by the choice between alternative classifiers can lead to different meanings. On the same page, Aikhenvald (*ibid.*) mentions the Australian language Murrinhpatha, where the stem *kamarl* glossed ‘eye’ is interpreted as ‘eye/face’ with the generic classifier *nanthi*, but ‘water-hole’ with *kura* (glossed cl:aquatic)’, ‘sweetheart’ with *kardu* (cl:human), and ‘seed’ with *mi* (cl:vegetable).

What all this shows is that structural hypotheses like the one sketched in (7) reflect the organization of basic abstract conceptual content, but their elements do not coincide with conceptual atoms and their combination does not reflect conceptual structure (see Ramchand and Svenonius 2014 and Acquaviva and Panagiotidis 2012).

*3.6 How language defines object concepts*

In section 3.3, we relied on the semantic notion of kind to model the content of object concept in language-internal terms. A few clarifications are necessary before proceeding.

 First, the kinds that correspond to an object concept are a subset of all the possible kinds. The reason is that, as explained in 3.3, in most treatments every property can be formally turned into a kind (an entity of type *e*) by abstracting intentionally over its instances across all possible worlds. This means that not only *dog* or *three-legged dog*, but also *dog with a particularly friendly character* or *dog with one ear up and the other down* could all underlie kinds. What we need is instead a kind that defines a type of object constant across contexts and indices. This does not correspond to the notion of “well-established kind” (Krifka 1995), which has practical utility but is too vague and open-ended to be of theoretical use. Instead, we make crucial reference to the rigidity of kinds as names for abstract individuals.

This move excludes properties presented as contingent and all references to an indexical centre, a speaker’s “here-and-now.” Therefore, it eliminates a whole class of what Carlson (1980) identified as ‘impossible kinds’ like *parts of this machine*, which simply points to unspecified entities anchored to the speaker by the indexical *this* and does not describe an object type independently of that context and what *this machine* refers to in it. A nominal can refer to an entity so described, as an object file (see 2.1 above). A simple noun, as opposed to a phrase, cannot generalize it as an object concept definable across possible worlds.

 Second, a grammatical structure does not have to correspond to a “minimal noun” to identify a kind and thereby an object type. The phrase *smelly garlic* describes an instance of the type *garlic* that is contingently smelly. Howoever, *wild garlic* describes a kind, and an object concept, even though it has the same adjective+noun structure. These considerations led Svenonius (2008) to identify a minimally noun-local domain hosting not only so-called noun classifiers (those most similar to gender in occurring independently of counting or other modification), but also non-intersective adjectives which help define a kind, as in *wild rice* or *high tea* (as opposed to *wild shout* or *high wall*). Not only gender, but also number has a role in defining certain kinds. Insofar as the “concrete stuff” reading of *brains* (also, extensively, “intelligence”) is distinct from the bounded single-organ interpretation of *brain*, it is the grammatical determination of plurality that expresses the difference. We conclude that there is no fixed-size piece of the nominal substructure that expresses an object concept.

 Third, object concepts may be instantiated by things taking place in time but this does not make them describable as eventive structures, like verbs. Hence, terms like ongoing process, culmination, or result state are not relevant for the description of their semantics, even for event-denoting nouns like *war*. These are conceptualized as continuants and not as occurrents, in classic philosophical terms (Simons 1987: 129–137), that is, as entities that are wholly there at each of different times, rather perduring through time so that only a (temporal) part of them is situated in a given time interval.

Thus, even if *day* denotes events extended in time, it can be true that the same day, as a whole object, is sunny at time *t1* and overcast at *t2.* This because the subject of predication remains the same while two contradictory predicates apply to it at different times. The nominal mode of denotation, then, differs from the verbal one not because nouns denote “things” rather than ‘”events” (and it seems hard to find a non-eventive denotation for *event*), but because it is characterized in ways that do not refer to event-structural properties like “completion” or “result state.” Correspondingly, no element in the structure in (7) makes specific reference to temporally extended part structure.

A final consideration must clarify the status of (nominal) lexical concepts, the object concepts lexicalized by nouns in the mental ontology of speakers. The existence of nouns like *lawyer* shows that the stock of nouns does not represent the inventory of what we believe to exist as objects. This holds not only because we speak about many discourse referents which we know do not exist[[5]](#footnote-6), but also because the very content of such concepts entails that they cannot denote independent entities but *phases* or *roles* of some entities. More subtly, we can also refer to entity types like *average American* and truthfully attribute properties to it which are true of no other entity (such as that of having 2.3 children). However, this in no way implies an ontological commitment to the existence of objects like abstract Americans along more familiar object concepts~~.~~

Therfore, as Collins (2017: 378) puts it, “semantics as actually pursued is an inquiry into structural conditions on the kind of thoughts we can express with linguistic forms independent of how the world actually is that would make our thoughts true or false.” There is strong evidence that language encapsulates as concepts objects that may only have a very indirect counterpart in our non-linguistic cognition (i.e. they are mind-internal, as argued in section 2). This does not mean that language can arbitrarily define no matter what entity. The properties of being mass or count, of being modifiable by dimensional adjectives, of belonging to a certain gender or noun class, or of being accompanied by certain classifiers, are not distributed at random over the lexicon of nouns. Language has its own constraints, and the linguistic encapsulation of concepts affords conceptualizations that do more than simply replicate sensory input or non-linguistic cognition.

*3.7 Possible objects, impossible nouns*

We saw in the preceding section that *parts of this machine* could refer to an object. However, the indexical *this* prevents it from identifying an object concept, as a generalization of possible referents across worlds independent of an indexical anchoring. This means that there cannot be a simple noun with that meaning. We claim that a simple noun expresses an object concept, and the grammatical context of a noun phrase turns that into a description for a particular object, corresponding to an object file in the visual system. Concepts can also be expressed by structures larger than a simple noun, in *high tea*, but simple nouns are a subset of concept-denoting expressions. Therefore, we exclude a whole range of meanings as possible contents for bare lexical nouns, namely all those expressible by a nominal phrase containing indexicals. This entails excluding the class of tropes, or particular instantiations of a property (like “the roundness of this ball”, “the height of Mary”), from possible noun contents (Moltmann 2009). Likewise, there can be no simple nouns with the values “segment from *a* to *b*” or “number of planets”, which imply a particular frame of reference (‘the point *here* marked as *a*’, “the planets in *our* world”).

 Consider now the notion of “thing seen”, i.e. having the property that there is at least some seer who at some time has seen it. It does not contain an indexical, and it is undoubtedly defined in visual terms. In addition, *thing* is by itself a noun, even though it is so radically underspecified that it only expresses information about countability. This is a problem for approaches that make countability dependent on perceptual cues, even non-deterministically, like Rijkhoff (2002). We do not therefore exclude this from the range of concepts potentially encapsulated by a noun. However, it would not be a visual object in our sense. This because it cannot correspond to any object concept determined visually based on a collection of stimuli, or the ‘visual types’ of section 2.1: being seen is simply not a visual type.

 This introduces a different sort of constraint. As discussed in 2.2, visual tokens must be organized together and recognized as instances of perceptual types, in order to be interpreted as object files. Minimal nouns, we assume, express object concepts, a notion which in the visual system corresponds to the generalization of an object file into a stable representation in long-term memory. However, this means that nouns cannot select an arbitrary subpart of the visual experiences that make up an object file and generalize a concept out of them. We will model more precisely in section 4 the relationship between the two representation systems. Here we note that this simple intuition suffices to exclude nouns for concepts like “every red patch in a picture” or “the rectangular shape of a book”.

 All of these notions can be thought, and can be expressed linguistically — but, we claim, by complex nominal structures only, not by simple nouns. The model of the language-vision relation presented in the following section deliberately disregards the language-internal distinction between nouns and phrases. This because what matters is the type-token relationship between elements of the representation, regardless of their “size”. It is therefore important to separate the predictions arrived at in the presente section from the conclusions to be reached in the next one.

**4 A hypothesis of the nouns-objects interface**

This section puts forward the main aspect of our proposal: a formal interpretation of the notions of visual and grammatical objects through which the two are related in a rigorous way. Such a programme, and its formal execution proposed here, rests on specific theoretical assumptions, which we spell out in 4.1. We then articulate the formal analysis in 4.2 in terms of an information system.

*4.1 The need for a meta-formalism to unify grammatical and visual representations*

We have presupposed that there is a clear distinction between language and object recognition as distinct cognitive modules. In this Section and Section 5, we propose that distinguishing these domains, and proposing specific constraints on how information is shared between them, makes specific predictions on what kind of possible nominals for visual objects can be found in a language. We thus discard most of the proposals falling within the “Construction Grammar” family (e.g. Goldberg 1995; Talmy 2000; Tomasello 2003; Goldberg 2005; Evans 2010). This because they assume that the cognitive mechanisms underpinning language and other domains of cognition, and by extension nouns and objects, are identical. Thus, these theories would predict that no mappings could arise between these two domains.

Other architectures envisage a distinction between cognition domains. Works such as Kirby (2012) suggest that non-linguistic systems like visual cognition or music can also be modelled via hierarchical structures. They thus display properties claimed to be exclusive to language (notably recursion: Hauser, Chomsky and Fitch 2002; Pinker and Jackendoff 2005). However, these architectures fall silent on the specific formal details by which these domains can be modelled, and on how distinct types of domain-specific information can be explicitly represented. Similarly, Jackendoff’s “parallel architecture” (Jackendoff 2002), proposes that cognition “modules” (in the sense of Fodor 1975, 1983) process different types of information in parallel. These different types can successively be synchronized or “bound”, by having the tokens of one information type (e.g. the object file for a blue ball) being bound with a token from another type (e.g. the NP *blue ball*).

A further alternative consists of those linguistic frameworks that explicitly propose an ontology for linguistic categories. Frameworks such as Sign-Based Construction Grammar (Sag et al. 2012) HPSG (Head-Driven Phrase Structure Grammar, Müller 2015), but also Auto-Lexical grammar (Saddock 2011) sketch the conditions by which linguistic expressions can be “connected” to non-linguistic representations. Similarly, the minimalist programme proposes that two central properties of language are recursion and the so-called “interface conditions”, conditions whereby linguistic expressions can be “connected” to non-linguistic information (Chomsky 1995, 2011; 2005a; Harley 2014). However, the details of these mappings to and from language representations are not worked out in detail in any of these theories. Therefore, a hypothesis of the relation between these two modules of cognition, beyond vague mentions of “interface conditions”, must be proposed.

For this purpose, we use Barwise and Seligman’s (1997) theory of information flow to connect a logic of visual objects with a logic of syntactic trees. The two apparently different frameworks can be connected via a presentation that highlights their formal, algebraic aspects. In this section, we first translate these logics into a format that is easily integrated with a logic of information flow, and then we discuss how we use this result to state precise predictions.[[6]](#footnote-7) We choose this theory as a meta-formalization over other theories of how nouns express entities (e.g. Gupta 1980; Murphy 2002; Lasersohn 2018) for three key reasons.

First, even in its most minimal presentation, it offers a simple yet elegant formalism that captures relations between encapsulated (“modular”) information systems and shows how these systems are connected. More philosophically oriented works (e.g. Gupta 1980; Lasersohn 2018) offer a greater focus on the properties of the nominal system, thereby not addressing what types may be at work in the visual system.

Second, the formal component allows us to make precise notions such as “classification”, “relation”, “type” and “token” (Barwise and Seligman 1997: Ch. 2–4; cf. also Smith 1995). Other works (e.g. Murphy 2002) propose that lexical content entirely reduces to conceptual knowledge. Thus, this formal system can also show how the nominal and the visual systems are nevertheless connected, although they are distinct, unlike other frameworks.

Third, via the notion of “infomorphism” and its application, the formal component provides inference rules on what counts as possible object and noun, and therefore as how these different notions constraint one another. Similar rules are presented in Gupta (1980). However, these rules can be said to only focus on how nominals are constrained in labelling concepts. As our discussion has suggested so far, hoeever, the two systems constrain one another at the more basic level of tokens and types. Thus, our hypothesis seems better to handle these data. We turn to its application.

*4.2 The information system*

An information system can be represented via a logical structure that captures the rules and the relations among the *types* of information that the system manipulates. In our case, one system manipulates visual information, and the other linguistic information. Two or more information systems can be connected via an *infomorphism*, defined as a mapping between two information systems that exchange information. The pre-theoretical intuition that governs this approach is that we study two systems as “parts” composing the “whole” of cognition. We then study how information is exchanged or *flows* between the two distributed systems. This flow of information can define constraints that are internal to each system (Barwise and Seligman 1997: Ch.1–3). For instance, visual objects cannot have certain properties, for the linguistic system would not be able to associate names to them.

The information system approach thus provides two types of inferencing rules. The first are rules that allow one to prove the formation of complex types via basic types and operators (e.g. from *a* and *b* we have *a⊓b*). The second are rules constraining the derivation of nominal expressions and visual objects from each other (i.e. from *a* a nominal type and *b* a visual type, we can have the mapping *f(a)=b*). Here, we spell out how these rules work in formal terms.

The two structures that we have defined for nominals and objects can be formalized as the *classifications* or *classifying structures* defined as follows:

(8) A *Visual Structure* (henceforth) *VO* is a sextuple *< O,\*O,⊨, ⊔,⊓,⊑ >* consisting of:

 - A set of tokens *O*

 - A set of types *\*O*

 - An instantiation relation *⊨*

 *-* The two operations *⊔* (join), *⊓* (meet)

 - The relation *⊑* (part-of)

(9) A *Nominal Structure* (henceforth) *NS* is a sextuple *< N,\*N,⊨, ⊔,⊓,⊑ >* consisting of:

 - A set of tokens *N*

 - A set of types *\*N*

 - An instantiation relation *⊨*

 *-* The two operations *⊔* (join), *⊓* (meet)

 - The relation *⊑* (part-of)

(10) Variables:

 - *o* is a member of *O* (a visual token)

 - \**o* is a member of \**O* (a visual type)

 - *n* is a member of *N* (a nominal token)

 - \**n* is a member of \**N* (a nominal type)

For both structures, the instantiation relation is represented as *a⊨b***:** a token *a* instantiates a type *b*, or alternatively *b* classifies *a*, whether visual objects or noun tokens are classified.[[7]](#footnote-8) We use this symbol to state that a token (visual, nominal) represents or stands for (i.e. it instantiates) a given type (Barwise and Seligman 1997: Ch.3). In Sections 2–3 we used the equivalent notation *type(token)* (e.g. *blue(x)*) to represent this relation for reasons of readability. Now that we have a full definition of this relation, we can use this equivalence (i.e. *type(token)* equivalent to *token⊨type*)to present our hypothesis in a streamlined manner.

Tokens and types have a precise definition within the infomorphism, quite different from the intuitive sense of tokens or types of production or processing. The tokens of the visual system/structure *VS* are the stimuli which, interpreted as instances of abstract types, define the conceptualizations corresponding to object files (see Section 2.1). For instance, a visual token interpreted as an instance of the visual types “red” and “ball” defines the object file of a red ball (cf. section 2). In the nominal system/structure *NS*, tokens are linguistic symbols considered as a choice of vocabulary material, independently of their grammatical content. These language-particular morphemes and lexical roots are sense types, but they are construed as tokens: they provide a lexical interpretation for a morphosyntactic template.

The types in *NS* are instead morphosyntactic representations, which express the grammatical information defining a nominal, but do not individuate any particular lexical choice. For instance, we can use the symbols *the* and *dog* to indicate a choice of vocabulary that lexically interprets the type defined by the morphosyntactic object *[Determiner [Noun]*. This because “Noun” is a shorthand for that deepest embedded substructure which is realized as the morphological word we call a noun (cf. section 3).

Tokens like *the* and *dog* in this perspective are not exponents, because not exponents but abstract symbols flesh out an abstract grammatical representation. Nor are they lexemes, or lexical items, or morphosyntactic words, because these notions presuppose a link between a choice of vocabulary item and its content, grammatical and non-grammatical. By relating these tokens to the type grammatical templates they lexically interpret, we define a particular construal of them, within our formal hypothesis, as lexical items abstracted away from their morphosyntactic value. They identify by a specific choice of linguistic symbols one token of a type of grammatical representation

The linguistic notion of a token-identifer for a mental representation provides a precise match for the “mental orthography” proposed by Margolis and Laurence (2007: 583) for the representation of concepts, as “the formal properties that allow the cognitive system to reidentify tokens of the same representation type”. In sum, a linguistic token of the nominal system *NS* is related to a grammatical template and thereby defines a particular noun, much as a visual stimulus is related to a visual type and thereby defines a particular object (in the sense of an object file, see Section 2.1).

The two operations *⊔* (join), *⊓* (meet) have been introduced in Section 2 to capture how visual types can be combined to form complex types (i.e. “prove” their existence). Recall that *a⊓b* (the meet of *a* and *b*) is obtained when a token instantiates two properties at once. For instance, *cone(x)⊓sphere(x)* is the complex type that a token instantiates when it instantiates an ice-cream object. In the nominal system, where types consist of abstract grammatical information, a meet type obtains when information about two grammatical features is lexically interpreted via a single token. This happens in so-called portmanteau affixation, where distinct morphological features are fused into a single morpheme. If we take the Italian name *gatt-a*, the number and gender inflection *–a* carries the *singular(x)⊓feminine(x)* features (short for the meet of *singular number* and *feminine gender*).

The join type can also represent a form of underspecification. If a ball’s colour is represented as *red(x)⊔blue(x)*, then it is still not clear whether the token of a ball instantiates either colour: a ball can be red or blue. This is how object files with a determinate choice of types can generalize to object concepts, like the concept of a ball with one of many possible colours (cf. 2.2). A parallel in the nominal system is provided by an English noun like *deer* being underspecified for number, since it can be used to lexically interpret a singular or a plural grammatical representation.

 The relation *⊑* (part-of) represents relations between types. Thus, *blue(x)⊑colour(x)* and *red(y)⊑colour(y)* represent the fact that the *blue* and *red* visual types are represented as “part” of the type domain of the *colour* super-type. A similar reasoning applies to nominal features: *male(z)⊑gender(z)* and *female(t)⊑gender(t)* represent the fact that these specific gender types are indeed part of the type domain of linguistic gender. As our discussion in Section 2–3 has shown, the introduction of these operations becomes necessary to represent the different notions of “complex type” and “type relation” that emerge from the analysis of visual and nominal domains. The introduction of these operations is also sufficient to represent these notions. Once we can represent how types are formed and related, we partially reconstruct notions such as mereology and connectedness (visual systems) or lexical relations (nominal systems). We also have a system that can prove the existence of complex types, objects, and nominals.

We now introduce the infomorphism as the second inference system. For this purpose, we use a graphical representation of the infomorphism, offered in Fig.1:

 *i*

 CONCEPT KIND

 *i*

 *File nominal*

 *g*

 \**O \*N*

 *⊨O ⊨N*

 *f*

 *O N*

**Figure 1.** The structure of our infomorphism.

Let us explain how to read the graphic in detail. First, the instantiation relation *⊨O* represents a relation holding between object tokens and types; the relation *⊨N* that betweennominal tokens and types. When a a visual stimulus/token instantiates a visual type, the instantation relation between the two entities holds, defining an object file. The upward arrows can be interpreted as representing this relation in short-term memory. A parallel process of instantiation holds between tokens in the nominal system and their types. A token such as *ball* can be assigned to a type defining a singular “noun” (a substructure of a noun phrase), resulting in the grammatical representation of an inflected noun: *[N:SING ball ]*.

For example, if the nominals (11a-b) are interpreted as the token roots in (12) instantiating the structure types in (13a-b), their semantic interpretation defines kinds. Not just what *lazy dog* or *the lazy dog* are true of in a given context and possibly world. As the term “kind” by itself allows various interpretations, especially in connection with its relation with object concepts, we adopt the neutral term *nominal kind*, which includes all kind-level intensional generalization of the content of a nominal, no matter what its size and its “established” nature. We notate it as in (14):

(11) a [Adj lazy [N dog ]] nominals

 b [ Det the [Adj lazy [N dog ]]]

(12) *lazy*, *dog* lexical choices (tokens in *NS*)

(13) a [Adj Adj [N Root ]]

 b [ Det Det [definite, sg] [Adj Adj [N Noun ]]] structures (types in *NS*)

(14) a [LAZY DOG]Kind nominal kinds

 b [THE LAZY DOG]Kind

This systematization does not identify a particular substructure of a nominal as the grammatical expression of an object concept, as we concluded in Section 3. Therefore, the information flow between the visual and linguistic systems does not reference grammatical representations of a specified size. There is no stipulation that object files and object concepts should correspond to certain substructures. The fact that a systematic connection only holds between between types and tokens means that visually- and grammatically-defined concepts are related, but indirectly. In both systems, the existence of such mental representations is conditional on a type being instantiated, or classified, by a token.

Let us now give the central definition. Given our classifications *VO* and *NS*, an infomorphism is defined as a pair of functions *i=*<*f*, *g>* such that:

(15) a *f(n)⊨\*O* if and only if *n⊨g(\*O)*

 b *f* establishes a mapping from *N* to *O*

c *g* establishes a mapping from *\*O* to *\*N*

In words, the classifications underpinning objects and nominals can be connected via two symmetrical functions (more accurately, mapping relations). First, the function *f* maps nominal tokens to their matching visual tokens (i.e. we have *f(n)=o* for any *n* and *o* belonging to *N* and *O*, respectively: Barwise and Seligman 1997: Ch. 3). Second, the function *g* maps visual types onto nominal types (i.e. we have *g(\*O)=\*N*, for any *\*n* and *\*o* belonging to *\*N* and *\*O*, respectively). An infomorphism *i* thus establishes mappings between two systems that manipulate different types and tokens of information: visual and linguistic.

The choice of this infomorphism is not innocent. The definition in (15) indirectly suggests that our mapping starts from the domain of lexical choices (the linguistic tokens in set *N*)to that of visual tokens (in set *O*) and sends information to the linguistic system, with the function *g* from visual types to nominal structures. There is an asymmetry: at the level of tokens, the nominal system carries information about the visual object system. A conjecture is that the system of lexical nouns (nominal substructures making up morphosyntactic words that head noun phrases) may constrain how the properties of visual objects can be realized in language. The constraints discussed in Section 3.7, which limit the range of possible concepts expressed by a simple noun, are an example of how properties of the linguistic system affect what can be perceived as an object (e.g. as a kind corresponding to an object concept).

The infomorphism is represented as a pair of diagonal arrows. The function *f* maps nominal tokens, members of the set *N*, onto visual tokens, members of the set *O*. Via the instantiation relation *⊨O*, nominal tokens become connected to visual types (cf. the diagonal mapping from *N* to *\*O*). In other words, an agent can connect (or “bind”) a visual stimulus to an inflected noun: one can be used to describe the other. The same holds for types, as the figure shows. Once an object file is accessed, the type of features it instantiates are connected to the types of features that the corresponding noun can instantiate, as the function *g* shows. The diagonal mapping between *O* and *\*N* establishes that the corresponding nominal is used to describe an object file *qua* a visual stimulus connected to a type. Thus, the fully flected noun *balls* can be used to represent several balls as visual objects, as we explain below.

Once object files and nominals are defined as distinct representations, their mapping becomes indirect, a fact we represent via the function *i*. An object file can be connected (or “bound”) to a nominal (i.e. we have *i(file)*=*nominal*), and so can a concept to a kind (i.e. we have *i(CONCEPT)*=*kind*).The relation between the tokens and types forming these higher-order mental representations is not accessible at these levels (i.e. no “diagonal” mappings are defined). In other words, an agent establishes that the noun *ball* describes an entity that belongs to the “ball” kind, and that an object file representing a ball belongs to the *BALL* concept. That object files and and nominals are bound becomes a possibility, rather than a consequence that can be inferred in context, as we clarify below.

Note, however, that the infomorphism is defined as working in both directions. Therefore, the properties of visual objects can indirectly constrain those of nominals, and *vice versa*. As our discussion suggests, these constraints pass through the notions of kinds and concepts, which we have shown to play a crucial role in both information systems. To fully appreciate the predictions that the hypothesis makes, however, we need to spell out how this infomorphism works, and how it connects nominals to visual objects.

**5 The Interface: Special and General Cases**

In our hypothesis of information flow, mappings work as *homomorphisms*, or “many-to-one” functions relating linguistic tokens to visual ones, and visual types to linguistic ones (Landman 1991: Ch. 2; Barwise and Seligman 1997: Ch. 2). This has notable consequences, as we discuss in this section on a case-by-case basis.

First, several tokens of the linguistic system instantiating the same nominal type can correspond to a single token of the visual system which instantiates a given type, given the mapping *f(n)=o*, where *n* is a variable over lexical material (noun tokens) and *o* a variable over visual tokens. Second, several visual types *\*o* can be mapped onto one linguistic type *\*n*, given the mapping *g(\*o)=\*n.* Thus, the function *f* can map mutually alternative lexical choices *n* for the same grammatical template *\*n* to the same visual token *o*. This accords with the fact that distinct lexical choices that are morphosyntactically equivalent (say, *ball* or *sphere* or simply *object*) may apply to the same visual experience, uniformly typed as the *sphere* visual type. Conversely, there are obviously different visual types that can be mapped to the same grammatical template, like the structure defining a noun phrase, or a smaller substructure, or even just what would be generally be called a noun.

These apparently simple observational facts reveal an important result. A formal representation of the relations between visual and nominal systems should at least be compatible with them. We therefore prove that the two functions *f* and *g* achieve this standard of empirical adequacy. A different way to relate the two systems, e.g. by establishing a direct correspondence between an object file (a visual token-type instantiation) and a lexically interpreted phrase (a linguistic token-type instantiation) would not derive this result.

The language-vision mapping is more indirect at the level of concepts and kinds, since we conceived concepts as unions of possible object files sharing types. Thus, to understand how concepts and kinds can be related, we must start by examining the token-type mappings (or “classification”) in each domain, then considering how they generalize at the higher order of nominals (lexically intepreted noun phrases), object files (map between token visual stimulus and type), kinds (generalizations of concepts), and concepts (generalizations of object files).

To best appreciate the (empirically correct) consequences of our approach, we first focus on cases where the infomorphism is defined over special elements: undefined and empty object tokens and types. We then turn to a brief illustration of standard values. Our method consists in first establishing, on the left side of a semi-colon, the token-type relation holding for each information system. Subsequently, we illustrate the overall relation between the visual and the linguistic domain which results from those mappings, given the definition of infomorphism *i* in (15).

*5.1. First Case: Undefined values*

An infomorphism may takes an undefined value, here represented as “⊥”, for either an object or a nominal token in (16). As an example, consider the constructed example *dax* (Landau et al. 1988, 1992), introduced as a label for an unspecified object with a novel shape for the observer. When speakers encounter such a novel noun, they will not be able establish what visual object this noun describes. The formalisms in (16) capture this form of undefinedness:

(16) a *(⊥⊨\*O)=⊥ , n⊨\*N; f(n)⊨\*O, n⊨g(\*O)*

 b *o⊨\*O, (⊥⊨\*N)=⊥; (f(⊥)⊨\*O)=⊥, g(\*O) = \*N*

 Consider (16a) first. If a visual token is undefined, it cannot instantiate any object type (that is, we have *(⊥⊨\*O)=⊥*). However, the function *f* can still map a linguistic token onto this undefined visual type (that is, *f(n)⊨\*O*). This means that a lexical choice *n* can map to a visual type *\*O* although there is no visual stimulus *o* that could instantiate this type. Because *\*O* is defined, *n* can be conceived as the token that instantiates a grammatical type resulting from the application of the function *g* to it (that is, *n⊨ g(\*O)*). Our hypothesis establishes that there can be nominal linguistic expressions, as relations between a lexical choice and a grammatical structure, with the function *i* mapping them onto undefined visual objects.

This result may look surprising at first sight. However, nouns and nominal structures void of conceptual content must be allowed, for a rather powerful reason. It must be possible that languages include novel nouns, i.e. nominals that describe objects that have not been encountered before, and must thus be evaluated as possible (or impossible) objects. This artificial scenario corresponds to the normal experience of encountering an unknown word, which is recognizable as a noun on grammatical grounds, but whose content remains to be filled in. Such “empty placeholders” must be possible, if only to account for the learning of new words, e.g. during language acquisition. There is another type of nominals that have been independently argued to fail to identify any kind. Forms like *contents*, *beginnings*, *furnishings* are true of objects in a particular (often contingent) spatiotemporal configuration or used for a particular function, but they do not describe what these objects are — even though the plural entails that they make up a collection (Wierzbicka 1988; Acquaviva 2008).

It is another question whether certain subtypes of undefined-concept nouns should be excluded on principled grounds. For instance, the imaginary *gavagai* which Quine (1960) constructed as a description for an undetached rabbit part is not void of conceptual content, no matter how problematic its underspecified nature could make communicating and using such a deficient characterization. It might be an impossible noun, since none of the infinite number of undetached parts of a rabbit (including the whole) could ever be unambiguously singled out for reference. However, its conceptual content, while probably useless, is not null, and our system does not rule it out.

Consider now (16b). In this case, what is undefined is a linguistic nominal token instantiating a nominal type (that is, we have *(⊥⊨\*N)*). Hence, we lack a nominal describing a specific object file (i.e. *(f(⊥)⊨\*O)=⊥*). However, a non-defined nominal token entails that no nominal type can be instantiated. Thus, no map between the two levels is defined (i.e., no expression). However, it is possible to have a token-type map in the visual system (*o⊨\*O*). A type *\*o* (in the set *\*O*)is defined, and the function g can define a linguistic type via such a visual type (*g(\*O)=\*N*). We therefore prove that a type of object can be conceptualized without a linguistic description for it, but with an appropriate grammatical categorization. An example can be a new biscuit “shape” that must still receive a (brand) name; other examples can be similarly defined. Our system entails that in such cases we would also “know” that the characteristic label for such a visual object would grammatically correspond to a nominal structure, by definition.

 We have considered either scenario where tokens are undefined, nominal (*n*) or visual (*o*). Undefined values can be assigned to visual and nominal types, too. The relevant formulae are in (17):

(17) a *(o⊨⊥)=⊥ , n⊨\*N; (f(n)⊨⊥)=⊥, (n⊨g(⊥))=⊥*

b *o⊨\*O, (n⊨⊥)=⊥; f(n)⊨\*O, (n⊨g(\*O))=⊥)*

The formula in (17a) proves that a visual token that does not instantiate any type and define an object file (i.e. *(o⊨⊥)=⊥*) cannot have a nominal to describe it. This because if a visual type *\*o* is lacking by hypothesis, *f* cannot define it on the basis of a linguistic token *n* (*f(n)⊨⊥)=⊥*); nor can *g* define it (*n⊨g(⊥))=⊥*). The entailment follows that no nominal can describe a visual token that has not been related to a type.

Correspondingly, (17b) proves that a nominal token not instantiating a nominal type (i.e. *(n⊨⊥)=⊥*) cannot serve as a basis for constructing a nominal phrase that describes a visual object. If the token-type instantiation on the linguistic side is undefined, the function *g* cannot relate a visual type *\*o* to a missing linguistic type (i.e. *(n⊨g(\*O))=⊥*). Our formal system thus derives in a principled way a very general and welcome result: the verbalization of a visual experience demands categorization. Given the definitions of *f* and *g*, a type is necessary for the two systems to dialogue, either a visual or a linguistic type. Uncategorized visual experiences cannot be verbalized except in cases like *that!*, a mere deictic which does not express *what* it is pointing to.

To see the relevance of this conclusion from the opposite, linguistic, perspective, consider a lexical choice that lacks the grammatical determination that would be necessary to fit it into a grammatical template. A case would be a noun appearing as an uninflected lexical root in a language that mandates inflectional morphology for that choice of noun. One example is the root *lamp-* instead of the inflected Italian word *lampo* ‘flash’. Its impossibility is strictly a language-internal matter, but we further derive the prediction that such an uninflected lexical choice cannot be related to a visual object.

*5.2. Second Case: Null values*

There are cases in which token and types are assigned a null value (represented as “∅”), a possibility which emerges from our use of Boolean algebras as representations of information systems. This value must be carefully distinguished from the undefined value. Null elements represent types or tokens that are part of a classification but express the “absence” of information. Undefined tokens and types instead are not *yet* part of a classification. For example, someone not yet acquainted with *red* as a visual type may associate a new, undefined type to a red object on first experiencing this colour. A null value, by contrast, corresponds to an absence of information. Hence, our system predicts that a class of nominals can name null or “empty” objects, e.g. *hole*.

To see what this means, consider first null object files and concepts, presented via the formulas in (18):

(18) a *∅⊨\*O, n⊨\*N; f(n)⊨\*O, n⊨g(\*O)*

 b *o⊨\*∅, n⊨\*N=NP; f(n)⊨\*∅, n⊨g(∅)*

The formula in (18a) proves that null visual tokens can instantiate visual types (i.e. *∅⊨\*O*), which can then be described by nominals, since nominal tokens and types can be mapped onto this null object token (i.e. we have *f(n)⊨\*O, n⊨g(\*O)*). The formula in (18b) proves that a visual type instantiating a null visual type (i.e. *o⊨\*∅*) can have a nominal kind to describe the object file.

These apparently paradoxical results have a simple interpretation. It is possible to have a null file defined via the meet of incompatible types. One example is *round(x)⊓square(x)*: *x* is a visual token that instantiates the meet type of round squares, corresponding to the object file *i* (i.e. we have *i=round(x)⊓square(x)*).[[8]](#footnote-9) Nothing bars the creation of a concept as a generalization from such a logically impossible (but not undefined) object file. Whether this could be the content of a simple noun, as opposed to a nominal phrase like *round square*, is another question (answered in the negative in \_\_\_\_\_\_). Concepts may also be null in a different sense: as conceptualizations for the absence of a shape, or a colour, and so on. Nominals such as *abyss*, *cavity*, *shadow* and similar others can describe such “null” concepts by their negation (cf. Casati and Varzi 1994). Our system thus predicts that visual objects can be also defined and identified via the absence of certain visual types (e.g. a hole in a doughnout as an absence of colour or texture). The system also predicts that there can be a class of nominals to precisely name these null objects.

Consider now the formulas in (19):

(19) a *o⊨\*O, ∅⊨\*N; f(∅)⊨\*O, ∅⊨g(\*O)*

b *o⊨\*O, n⊨\*∅; f(n)⊨\*O, n⊨g(\*O )*

According to (19a), a nominal type with a null token instantiating it (that is *∅⊨\*N*) can still be used to describe an object type and a corresponding token (so, we have *f(∅)⊨\*O, ∅⊨g(\*O)*). The formula in (19b) says that a nominal token may instantiate a null type (so that *n⊨\*∅*) but may nevertheless describe a specific object type (that is, *f(n)⊨\*O, n⊨g(\*O)*). The interpretation of these formulae presents some interesting results.

In the first case, a language may lack a word for an object, which can nevertheless be described via a certain nominal. For instance, English lacks a word for *maltagliati*, a specific pasta shape from central Italy. However, the word can appear in an *English* nominal, as a word that by assumption is not part of the English vocabulary, to describe this object concept. *Maltagliati* may still not be an English noun, but it can be used in English sentences to describe a certain visual object.

In the second case, in (19b), a non-null linguistic token instantiates a null nominal type, which despite being null can be related to a non-null conceptual representation. The justification for this kind of scenario is best seen considering types that correspond to grammatical categories. Languages with a two-value gender system may have nominals expressly marked as taking neither value (as opposed to being compatible with either one). Such is the case of “neuter” determiners like the Spanish *lo* ‘the’ or pronouns like the Spanish *esto / eso / aquello* ‘this / that’ (Ledgeway 2012: 106).

Regardless of the theoretical analysis, our system affords this descriptive interpretation. A specific morphological marking on these nominals (nominal token) determines a particular conceptualization (object concept), characterizing these forms as outside the gender system which otherwise underlies all nominals in the language. Similarly, a nominal form may determine an interpretation falling outside the values defined by a language number system. This happens with forms expressing a “general” value, compatible with singular or plural interpretation but distinct from either one (e.g. Corbett 2000, Paul 2012).

The following Malagasy examples (from Paul 2012: 103) show that the bare nominal *vorona* is not ambiguous between a singular and a plural reading, as it can license as continuation either a conjunction or a disjunction, unlike where a singular reading is made explicit (pst and at indicate ‘past’ and ‘actor topic’ in Paul’s glosses):

(20) a *Nahita vorona aho — ganagana na fony / ganagana sy fony.*

 pst.at.see bird 1.sg — duck or pigeon / duck and pigeon

 ‘I saw one or more birds — a duck or a pigeon / a duck and a pigeon.’

 b *Nahita vorona iray aho — ganagana na fony / # ganagana sy fony.*

 pst.at.see bird one 1.sg — duck or pigeon / duck and pigeon

 ‘I saw one bird — a duck or a pigeon / a duck and a pigeon.’

Proper names offer a subtler illustration. The literature on names has traditionally considered almost only personal names (see Matushansky 2008 and Predelli 2017 for a sample of opposing views). However, place names have grammatical properties setting them apart as a distinct class (cf. Moltmann 2015). One of them is the lack of so-called “expletive” definite articles, which some languages can attach to a name without for that turning it into a common noun. For example, dialectal German features both *der Hans* and *Hans*, but a place name like *Berlin* can never become \**der/das Berlin*. An article appears with a place name either always, as in *Den Haag*, or never. Insofar as place and personal names are the same grammatical objects, yet have distinct linguistic properties, there is no distinct “place name” nominal type. Complex place names like *The North Sea* (cf. Langendonk 2007: Ch. 4), but also *Isle of Man*, *Stratford-upon-Avon* suggest a similar conclusion. They partially describe the named place, but would hardly qualify as definite adjectival phrases with descriptive content. Their syntactic structure does not express their content as place names.

The upshot of our discussion is that the infomorphism can already account for several patterns regarding the relation between visual objects and nominal kinds under a unified perspective. This is possible even if this relation is indirectly defined via the two mappings forming *i*,those between tokens and types of each domain. Therefore, the infomorphism can already proves the existence of at least four types of nominals: novel nouns (e.g. *dax*), nouns for null objects (e.g. *shadow*)*,* “general” nouns (e.g. *varana*), and proper names (e.g. *Hans*, *Berlin*). The next section addresses the nominal types emerging from standard values.

*5.3. Third Case: Standard values*

Standard values involve join, meet and relational visual types to define complex object files. Nominals can then describe these complex types, but without stating which basic types form each complex type. This fact holds whether a given nominal is assigned a unique type (e.g. *singular*), or a complex type (e.g. *singular, feminine*). We thus focus on the full expressive power of the first inference system constituting the information flow theory. The pre-theoretical intuition on how this system works can be defined as follows.

Visual and nominal systems, as independent but parallel systems, prove how we can recognize complex object types as combinations of simple object types (e.g. from *cone* and *sphere* to *ice-cream*). They also prove how we can recognize nominal *gatt-o* as the combination of the root *gatt-* and the inflectional morpheme *-o*. When these systems are connected and exchange information involving non-null tokens and types, new constraints emerge on this exchange. We first discuss single-type cases via the formulae in (21), to show this fact:

(21) a *o⊨(a⊔b), n⊨N, f(n)⊨(a⊔b), n⊨g(a⊔b)*

 b *o⊨(a⊓b), n⊨N, f(n)⊨(a⊓b), n⊨g(⊓b)*

 c *o⊨(a⊑b), n⊨N, f(n)⊨(a⊑b), n⊨g(a⊑b)*

The formula in in (21a) says that visual tokens instantiating join visual types (i.e. *o⊨(a⊔b)*) define standard object files. These files can have nominals that describe these visual types (i.e. we have *f(n)⊨(a⊔b), n⊨g(a⊔b)*). Visual underspecification/ambiguity corresponds to the definition of a concept that can have a “general” nature. For instance, a nominal such as *colour* can describe a concept for which object files can be red, or blue, or yellow, or any other colour.

The formula in (21b) states that object tokens instantiating meet types (i.e. *(a⊓b)*) can have nominals describing these token-type pairs (i.e. *f(n)⊨(a⊓b), n⊨g(⊓b)*). From the definition of such object files and the possibility to describe these files, we can infer that nominals can also describe their corresponding concepts. Our hypothesis thus predicts that nominals such as *beagle* and *greyhound* can name specific *DOG* concepts, even if they may not explicitly represent what visual types make up these concepts. For instance, a visual stimulus *x* can correspond to a greyhound when it instantiates the meet of visual types for greyhounds (e.g. *four-legged(x)⊓ animate(x)⊓gray-mantle(x)*)*.* The nominal *greyhound* can be used to describe a specific “greyhound” object file. Hence, it acts as a label for the combination of these types into a single object file.

Thus, our system can also prove that nominals name objects that can be defined via the combination of more specific properties. At the same time, it predicts that this mapping is not isomorphic, i.e. a one-to-one correspondence. While *greyhound* is a singular noun, the objects it can describe need not to come in singular units, as generic statements show (e.g. *the greyhound is a beautiful canine species*: Carlson 1980; Chierchia 1998).

The formula in (21c) says that visual tokens instantiating types that are part of other types (i.e. *o⊨(a⊑b)*) can have nominals that describe these relations (i.e. *f(n)⊨(a⊑b), n⊨g(a⊑b)*). Our hypothesis thus predicts the existence of relational nominals such as *front*, *axis,* and several others, possibly in combination with prepositions like the English *of*. The central characteristic of these nouns can be described as follows. An observer can identify an object as part of a larger whole: a surface including a door can be identified as the “front” of a house. The nominal *front* can thus used to describe not only this object, but also the mereological relation to the object described by the nominal *house* (cf. Cruse 1986; 2004). More in general, our system predicts the existence of relational nouns as nominals that “embed” information about the relations that the objects they name have with other objects.

Overall, our system predicts that complex types exist but also that they can stand in logical (i.e. part-of) relations. It also predicts that the infomorphism can represent these relations in a systematic manner. Thus, our information system predicts what happens when the two systems exchange empty or undefined information, but also when they exchange “standard” information. With this result, we have built enough evidence to show that our account can cover a wealth of data regarding the relation between nouns and objects. Crucially, this account is possible because it is rooted in a mind-internal perspective about objects (cf. Pylyshyn 1988; Scholl 2001), and nominals (cf. Krifka 1998; Collins 2017). If objects are abstract categorizations for visual inputs, and nominals act as labels for these categorizations, then one would expect that nominals/labels for undefined or “null” objects could exist. Given how these systems exchange information, it is not necessary that this information is factual, for it can simply represent how we *could* name objects or discuss their paradoxical properties.

**6 Conclusion**

Through language, we elaborate a notion of object that corresponds to the meaning associated with a certain kind of expressions. The question we have considered is how these notions are related. The question, evidently, is too generic to allow a meaningful answer. We have proposed a way to make it more amenable to investigation and to empirical verification, by making precise in formal terms the notions of “object for visual cognition” and “object for linguistic cognition”. The literature offers many accounts of what it means to conceptualize an object in vision and in language. We have based our proposal on two independently justified approaches, which were then formally reinterpreted in terms of token-type instantiation relations. This allowed us to model their interaction by means of an infomorphism, constituted of two functions that relate representations in the visual and linguistic system to each other.

The significance of this exercise is that it generates predictions. A naive conception of the pairing of visual and linguistic objects might limit itself to considering concrete visible entities, in their cognitive-visual and linguistic encapsulation. Instead, our model by its own logic predicts as a necessary consequence a range of language-vision mappings even where a token (lexical material or visual stimulus) is undefined. More importantly, however, is that some mappings are predicted *not* to happen. Some notions like “parts of this machine” or “round square”, we argue, can be described by complex nominals but not by nouns because they cannot correspond to a suitably generalizable concept (section 3.7).

Far more important is the prediction that no nominal expressions, in any natural language, can describe bare stimuli not categorized as tokens of a type, and that bare linguistic material not categorized as spelling out a grammatical structure cannot describe a visual experience. A linguistic expression cannot linguistically describe a visual experience, unless this has been categorized (relating visual tokens to visual types). A visual experience cannot be referred to by a simple choice of words or morphemes, unless these express a specific grammatical structure (section 5.1). These are substantive falsifiable results, not just because our account rules out mappings that would be thinkable a priori, but also because they give us the theoretical vocabulary to describe these attested and purely hypothetical mappings. In particular, they allow us to distinguish undefined from null values for tokens or types, a distinction that captures cases of grammatical or conceptual under-determination (section 5.2). All of this, we submit, would be impossible if vision and language were compared just by pairing visually describable objects to grammatically well-formed nominals.

However, we think that our most important result is that of having made the question possible in the first place. “Thing seen”, “thing thought”, and “thing denoted by a noun” are impossibly vague notions. Therefore, we have relied on a large body of literature to extract a clear sense from them. We have spelled out a formal apparatus to represent them and, above all, their interaction. It is for this reason that the infomorphism defined in section 4 lies at the heart of our proposal. As a higher-level formalization, it allows the formal models of visual objects and of nominal grammatical objects to communicate with each other, making them commensurable and leading to specific, empirically verifiable conclusions. A more adequate analysis of visual objects and of nominals is evidently an open challenge. Still, over and above our contribution to the understanding of these two aspects of cognition, our key goal has been to indicate a rigorous way to relate them.

**Acknowledgements**

This research was supported by University College Dublin Seed Funding Grant 45263 (Paolo Acquaviva) ‘A theory of possible lexical items for visual objects’, and by grant 09-RNP-089 of NetWordS - European Network on Word Structure: Cross-disciplinary Approaches to Understanding Word Structure in the Languages of Europe' (Ursini). This funding facilited travel and made possible discussion and the analysis here reported. We gratefully acknowledge the support received from the two agencies. The authors wish to thank the editors for their editorial support, and two anonymous reviewers for the thorough feedback. The usual disclaimer applies.

**References**

Acquaviva, Paolo, and Phoevos Panagiotidis. 2012. Lexical decomposition meets conceptual atomism. *Lingue e linguaggio* XI: 165–180.

Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford Universit Press.

Acquaviva, Paolo. 2018. Names as kinds: An analysis of naming verbs. Ms, University College Dublin.

Adger, David. 2013. *A syntax of Substance*. Cambridge, MA: MIT Press.

Aikhenvald, Alexandra. 2003. 2000 *Classifiers. A Typology of Noun Categorization Devices.* Oxford: Oxford University Press.

Alvarez Fernandez, Susana and Maria Vanrell. 2011. Texton hypothesis revisited: A bag-of-words approach to combine textons. *Pattern Recognition* 45 (12): 4312-4325.

Anderson Stephen J., Noriko Yamagishi, Vivian Karavia. 2002. Attentional processes link perception and action. *Proceedings Royal Society London Biological Sciences* 269: 1225–1232

Anderson, Curt, and Marcin Morzycki 2015. Degrees as kinds. *Natural Language and Linguistic Theory* 33/3:791–828.

Asher, Nicholas. 2011. *Lexical Meaning in Context: A Web of Word*s. Cambridge: Cambridge University Press.

Baker, Mark. 2003. *Lexical Categories*. Cambridge: Cambridge University Press.

Barwise, Jon and Jerry Seligman. 1997. *Information Flow: The Logic of Distributed Systems*. Cambridge: Cambridge University Press.

Ben-Shahar, Ohad, Brian J. Scholl and Steven W. Zucker. 2007. Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research* 47 (9): 845–860.

Biederman, Irwin.. 1987. Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review* 94(1): 115–147.

Binnick, Robert (ed.). 2016. *The Oxford Handbook of Tense and Aspect*. Oxford: Oxford University Press.

Borer, Hagit. 2005a. *In Name Only*. Vol. 1 of *Structuring Sense*. Oxford: Oxford University Press.

Borer, Hagit. 2005b. *The Normal Course of Events*. Vol. 2 of *Structuring Sense*. Oxford: Oxford University Press.

Carlson, Gregory. 1980. *Reference to Kinds in English*. New York: Garland.

Casati, Roberto and Achille Varzi. 1994. *Holes and other superficialities*. Cambridge, MA: MIT Press.

Chierchia, Gennaro. 1998. Plurality of Mass Nouns and the Notion of Semantic Parameter. In Susan Rothstein, (ed.), *Events and Grammar,* Dordrecht: Kluwer, 53–103.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Collins, John. 2017. The Semantics and Ontology of The Average American. Journal of Semantics 34/3: 373–405.

Cooper, Robin. 2005. Records and Record Types in Semantic Theory. *Journal of Logic and Computation* 15 (2): 99–112.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Cowan, Nelson. 1995. *Attention and Memory: An Integrated Framework*. Oxford: Oxford University Press.

Cowan, Nelson. 2005. *Working Memory Capacity*. Hove: Psychology Press.

Croft, William. 2000 Parts of speech as typological universals and as language particular categories, in Petra Maria Vogel and Bernard Comrie(eds), *Approaches to the typology of word classes,* 65-102. Berlin: Mouton de Gruyter.

Cruse, Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Cruse, Alan. 2004. *Meaning in language: An introduction to semantics and pragmatics*. Cambridge: Cambridge University Press.

Cumming, Samuel. 2008. Variabilism. *Philosophical Review* 117.4:525–554.

Cyrino, Sonya, and Maria Teresa Espinal. 2015 Bare nominals in Brazilian Portuguese: More on the DP/NP analysis. *Natural Language and Linguistic Theory* 33/2: 471–522.

Déchaine, Rose-Marie, C. Mudzingwa, MartinaWiltschko and R. Girard. 2014. The internal syntax of Shona Noun Class Prefixes. *Language sciences 43:* 18-46 (special issue on “Exploring grammatical gender”).

Embick, David, and Alec Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39:1–53.

Evans, Vyvian. 2010. Figurative language understanding in LCCM Theory. *Cognitive Linguistics.* 21(4): 601–662.

Fábregas, Antonio, and Sergio Scalise. 2012. *Morphology: From Data to Theories*. Edinburgh: Edimburgh University Press.

Fassi-Fehri, Abdelkader. 2018. *Constructing the Feminine to Mean: Gender, Number, Numeral, and Quantifier extensions in Arabic*. New York: Lexington Books.

Feldman, Jacob. 2007. Formation of visual “objects” in the early computation of spatial relations *Perception and Psychophysics* 69 (5): 816-827.

Fodor, Jerry A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, Jerry A. 1983. *Modularity of Mind*. Cambridge, MA: MIT Press.

Gao, Tao, Gregory E. Newman and Brian J. Scholl. 2010. The Wolfpack Effect: Perception of Animacy Irresistibly Influences Interactive Behavior. Psychological Science 21(12): 1845–1853.

Geach, Peter. 1962. *Reference and Generality*. Ithaca, New York: Cornell University Press.

Gibson, James J. 1966. *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Ginzburg, Johnathan. 2005. Situation Semantics: The Ontological Balance Sheet. *Research Language Computation* 3 (2–3), 363–389.

Goldberg, Adele. 1995. Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Griffin, Nicholas. 1977. *Relative Identity*. Oxford: Clarendon Press.

Gupta, Anil. 1980. *The Logic of Common Nouns*. New Haven: Yale University Press.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90: 185–229.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225–276. DOI: <https://doi.org/10.1515/tl-2014-0010>.

Hauser, Marc D., Noam Chomsky and William Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569–1579.

Hummel, John I. and Irvin Biederman. 1992. Dynamic Binding in a Neural Network for Shape Recognition. *Psychological Review* 99 (3), 480-517.

Jackendoff, Ray and Steven Pinker. 2005 The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition* 97: 211–225.

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

Jeshion, Robin. 2015. Referentialism and predicativism about proper names. *Erkenntnis* 80:363–404.

[Julesz, B](https://en.wikipedia.org/wiki/Bela_Julesz)ela. 1981. Textons, the Elements of Texture Perception, and their Interactions. *Nature*. 290 (5802): 91–97.

Kahneman, D., Treisman, A., and Gibbs, B.J. 1992. The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24: 175–219.

Kirby, Simon. 2007. The evolution of language. In Richard Dunbar, R. and Louis Barrett, (eds.), *The Oxford Handbook of Evolutionary Psychology*: 669–681.

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In In G. Carlson and F.-J. Pelletier (eds), *The Generic Book*, University Of Chicago Press, 398–411.

Krifka, Manfred. 1998. The origins of telicity. In S. Rothstein (ed.), *Events and Grammar*, Dordrecht: Kluwer, 197-235.

Krifka, Manfred. 2004. Bare NPs: Kind-referring, Indefinites, both, or neither? In R. B. Young and Y. Zhou (eds), *Proceedings of Semantics and Linguistic Theory (SALT) 13*, University of Washington, Seattle. Cornell: CLC Publications.

Kripke, Saul. 1972. *Naming and Necessity.* Cambridge, MA: Harvard University Press.

Landman, Fred. 1991. *Structures for Semantics*. Dordrecht: Kluwer.

Langendonck, Willy van. *Theory and Typology of Proper Names*. Berlin and New York: Mouton de Gruyter, 2007.

Landau, Barbara; Linda B. Smith, L.B.; Susan Jones. 1988. The importance of shape in early lexical learning. *Cognitive Development*. **3**: 299–321.

Landau, Barbara, Linda B. Smith & Susan Jones. 1992. Syntactic context and the shape bias in children’s and adults’ lexical learning. *Journal of Memory and Language* 31: 801–825.

Lasersohn, Peter. 2018. Common Nouns as Variables: Evidence from Conserativity and the Temperature Paradox. In Rob Truswell, Chris D. Cummins, Caroline Heycock, Brian Rabern and Hannah Rohde (eds.). *Proceedings of Sinn und Bedeteung* 21(2): 731–746.

Laurence, Stephen, and Eric Margolis. 1999. Concepts and cognitive science. In Eric Margolis and Stephen Laurence (eds), *Concepts: Core Readings*, 3-82. Cambridge, MA: MIT Press.

Ledgeway, Adam. 2012. *From Latin to Romance*.Oxford: Oxford University Press.

Link, Godehard. 1983. The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretic Approach. In Rainer Bäuerle, Christoph Schwarze and Arnim von Stechow (eds), *Meaning, Use and Interpretation of Language*. Berlin: de Gruyter, 303-323.

Margolis, Eric and Stephen Laurence. 2007. The ontology of concepts—abstract objects or mental representations? *Noûs* 41: 561-593.

Marr, David. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Freeman and Co., SanFrancisco, California.

Massam, Diane (ed.). 2012. *Count and Mass Across Languages*. Oxford: Oxford University Press.

Matushansky, Ora. 2008. On the linguistic complexity of proper names. *Linguistics and Philosophy* 21:573–627.

Michotte, Albert. 1963. *The perception of causality*. London: Methuen.

Mitroff, Stephen R., Brian J. Scholl, and Karen Wynn. 2005. The relationship between object files and conscious perception. *Cognition* 96 (1): 67–92.

Moltmann, Friederike. 2009. Degree structure as trope structure: a trope-based analysis of positive and comparative adjectives. *Linguistics and Philosophy* 32: 51–94.

Moltmann, Friederike. 2015. Names and the mass-count distinction. Ms., CNRS/Université Paris 1.

Mueller-Reichau, Olav. 2006. Sorting the world. On the relevance of the kind-level / object-level distinction to referential semantics. PhD Dissertation, Universität Leipzig.

Müller, Stefan. 2015. *Grammatical theory: from Transformational Grammar to contstraint-based approaches*. Berlin: Language Science Press.

Murphy, Gregory L. 2002. *The big book of concepts.*Cambridge, MA: MIT Press.

Murphy, Gregory L., & Brian Ross, 2010. Category vs. object knowledge in category-based induction. Journal of Memory and Language, 63, 1-17.

Parsons, Terence. 1980. *Non-existent objects*. New Haven: Yale University Press.

Panagiotidis, Phoevos. 2014. *Categorial Features: A Generative* *Theory of Word Class Categories.* Cambridge: Cambridge University Press.

Paul, Ileana. 2012. General number and the structure of DP. In Massam (ed.), 2012, 99–111.

Pietroski, Paul, Terence Lidz, Timothy Hunter and Justin Halberda. 2009. The meaning of ‘Most’: semantics, numerosity and psychology. *Mind and Language* 24(5): 554–585.

Predelli, Stefano. 2017. *Proper Names. A Millian Account*. Oxford: Oxford University Press.

Pylyshyn, Zenon. 1984. *Comptuation and Cognition: Towards a Foundation for Cognitive Science.* Cambridge, MA: the MIT Press.

Pylyshyn, Z. 1989. The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition* 32 (1): 65-97.

Pylyshyn, Zenon. 2004. Some puzzling findings in object tracking, I: tracking without keeping track of object identities. *Visual Cognition* 11 (7): 809–822.

Pylyshyn, Zenon. 2006. Some puzzling findings in multiple object tracking (MOT): II. Inhibitation of moving nontargets. *Visiaul Cognition* 14 (2): 175–198.

Quine, Willard van Orman. 1960. *Word and Object*.Cambridge, MA: MIT Press.

Ramchand, Gillian, and Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46: 152–174.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon*. Cambridge: Cambridge University Press.

Riesenhuber, Max and Tommaso Poggio. 1999a. Hierarchical models of object recognition in the Cortex. *Nature Neuroscience* 2(8): 1019–1025

Riesenhuber, Max and Tommaso Poggio. 2000. Models of object recognition. *Nature Neuroscience* 3(8), 1199-1204.

Riesenhuber, Maximilian and Tomaso Poggio. 2002. Neural mechanisms of object recognition. *Current Opinion in Neurobiology* 12: 162–168.

Rijkhoff, Jan. 2002. *The Noun Phrase*. Oxford: Oxford University Press.

Rothstein, Susan. 2010. Counting and the mass-count distinction. *Journal of Semantics*, 27, 343–397.

Rothstein, Susan. 2004. *Structuring Events*.Oxford: Blackwell.

Saddock, Jerrold M. 2012. *The Modular Architecture of Grammar*. Cambridge: Cambridge University Press.

Sag, Ivan A. Hans C. Boas, and Paul Kay. 2012. Sign-Based Construction grammar: an informal synopsis. In Boas, Hans C. and Ivan A. Sag (eds), *Sign–Based Construction Grammar*. Stanford, CA: CSLI. 69–189.

Scholl, Brian. 2001. Objects and Attention: the State of the Art. *Cognition* 80 (1): 1–46.

Scholl, Brian. 2007. Object persistence in philosophy and psychology. *Mind and Language*, 22(5): 563-591.

Scholl, Brian. J. 2009. What have we learned about attention from multiple object tracking (and vice versa)?. In D. Dedrick and L. Trick (eds), *Computation, cognition, and Pylyshyn*,49–78. Cambridge, MA: MIT Press

Scholl, Brian. J., and Zenon W. Pylyshyn 1999. Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology* 38: 259–290.

Scholl, Brian J., and Paul Tremoulet. 2000. Perceptual causality and animacy. *Trends in Cognitive Sciences* 4: 299–309.

Scholl, Brian and Tao Gao. 2013. Perceiving Animacy and Intentionality: Visual Processing or Higher-Level Judgment? In Marcus D. Rutherford and Val A. Kuhlmeier (eds), *Social perception: Detection and interpretation of animacy, agency, and intention,* 197-230. Cambridge, MA: MIT Press.

Scholl, Brian, Zenon W. Pylyshyn and Jacob Feldman. 2001. What is a visual object? Evidence from target merging in multiple object tracking. *Cognition* 80 (1): 159–177.

Shuwairi, Sarah M., Bainbridge, Rebecca, & Murphy, Gregory. L. 2014. [Concept formation and categorization of complex, asymmetric, and impossible figures](https://psych.nyu.edu/murphy/ShuwairiEtAl.14.pdf). *Attention, Perception, & Psychophysics, 76*, 1789-1802.

Simons, Peter. 1987. *Parts. A Study in Ontology*. Oxford: Oxford University Press.

Smith, Barry. 1995. Formal Ontology, Common Sense and Cognitive Science. *International Journal of Human-Computer Studies,* 41: 641–66.

Strawson, Peter. 1959. *Individuals. An essay in descriptive metaphysics*. Oxford: Routledge.

Svenonius, Peter. 2008. The position of adjectives and other phrasal modifiers in the decomposition of DP. In Louise McNally and Christopher Kennedy (eds), *Adjectives and Adverbs,* Oxford: Oxford University Press, 16–42.

Talmy, Leonard. 2000. *Towards a Cognitive Linguistics*. Cambridge, MA: MIT Press.

Tomasello, M. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Tremoulet, Paul D., and John Feldman. 2000. Perception of animacy from the motion of a single object. *Perception*, 29: 943–951.

Ullman, Shimon. 1979. *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

Ullman, Shimon. 1996. *High Level Vision*. Cambridge, MA: MIT Press.

Ursini, Francesco-Alessio. 2016. Objects and Nouns: An infomorphic theory of the Vision-Language Interface. In Jordan Zlatev, Göran Sonesson and Piotr Konderak (eds.), *Meaning, Mind and Communication: Explorations in Cognitive Semiotics*, 330–342. Frankfurt: Peter Lang.

Ursini, Francesco-Alessio. 2011. Space and the Vision-Language Interface: A Model-Theoretic Approach. *Biolinguistics* 5 (3): 170–225.

Van Marle, Kristy and Brian J. Scholl. 2003. Attentive tracking of Objects versus Substances. *Cognitive Science* 14 (5): 497–504.

Wells, Andrew J. 2002. Gibson's affordances and Turing's theory of computation. Ecological psychology, 14 (3): 140-180

Whitagen Randy, Hermann de Poel, Daniel Araújo, Gerry-J. Pepping. 2012. Affordances can invite behavior: reconsidering the relationship between affordances and agency. *New Ideas Psychol*. 30: 250–258.

Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam-Philadelphia: John Benjamins.

Wiggins, David. 1980. *Sameness and Substance*. Oxford: Blackwell.

Wiltschko, Martina. 2012. Decomposing the Mass/Count distinction: Evidence from languages that lack it. In Massam (ed.) 146–171.

Wolfe, Jeremy M. and Sara C. Bennett. 1997. Preattentive Object Files: Shapeless Bundles of Basic Features. *Vision Research* 37 (1): 25–43.

Zacks, Jeffrey M. 2004. Using movement and intentions to understand simple events. *Cognitive Science* 28: 979–1008.

Zacks, Jeffrey M. and Kehna M. Swallow. 2007. Event Segmentation. *Current Directions in Psychological Science* 16: 80–84.

Zamparelli, Roberto. 2000. *Layers in the Determiner Phrase*. New York: Garland.

Zhang, Niina. 2012. Countability and numeral classifiers in Mandarin Chinese. In Massam (ed.) 2012, 220–237.

Zhang, Niina. 2013. *Classifier Structures in Mandarin Chinese*.Berlin: De Gruyter.

Zhu, Songchun, Chengen Guo, Yizhou Wang and Zijan Xu. 2005. What are textons? *International Journal of Computer Vision* 62(1): 121–12.

1. Affordances and motion typs are sometimes conceived as more “intensional”, or “higher order” visual types (cf. Wells 2002; Scholl and Gao 2013). In this article, we focus on these visual types as the starting points for such inferences, and leave the discussion of such inferences on non-visual information for future research. [↑](#footnote-ref-2)
2. Meet and join are *associative*, *idempotent* and *commutative* operations. If an object is blue ball, then the order by which we individuate these types (commutativity), their possible combinations (associativity) and identity (idempotence) do not matter. These aspects are not crucial to our discussion. [↑](#footnote-ref-3)
3. See Baker (2003) and Panagiotidis (2014) for recent formal approaches to the bases of nominality. In philosophy, the definition of nameable entity is a topic rooted in debates on identity (Geach 1962; Griffin 1977; Gupta 1980; Wiggins 1980). Lasersohn’s (2018) proposal that nouns are variables for individuals is consonant with our approach, as is Cumming’s (2008) analysis of proper names as free variables. [↑](#footnote-ref-4)
4. The notion of possible world finds application in formal semantics to state truth conditions of modal sentences: *p* is possible if there exist alternative worlds where *p* holds. In our context, we could imagine that some objects are depicted as having different properties from the real ones, e.g. in the fictional worlds of comics. However, a noun such as *ball* used in a comic rigidly refers to the same *BALL* concept labelled by the noun in the real world, even though the objects falling under it may be slightly different. [↑](#footnote-ref-5)
5. The literature on non-existent objects contains contributions showing that an internalist approach is best equipped for accounting this range of empirical challenges. See Parsons (1980); Jeshion (2015); Collins (2017). [↑](#footnote-ref-6)
6. Note that Barwise and Seligman (1997) presents information theory as mostly an externalist view about systems. However, the nature of the represented system determines the overall view (e.g. switchboxes, in their formulation), as the authors acknowledge (Barwise and Seligman 1997: Ch. 2). Thus one can use the theory to show how a system can exchange information with the environment (externalist view), or how a system can represent and store information (internalist view), or even integrate the two information flows into an extended view (cf. Ursini.2011, 2016). We thank an anonymous reviewer for raising this issue. [↑](#footnote-ref-7)
7. Let us note that the symbol “*⊨*” is used here with a different interpretation from propositional logic, in which it represents an entailment relation (cf. Landman 1991: Ch. 1). [↑](#footnote-ref-8)
8. In set-theoretic terms, we can say that the intersection of the set of round objects and square objects contains only the empty set, with (visual) tokens interpreted as elements of sets/types. [↑](#footnote-ref-9)