



HAL
open science

Boundaries in space and time: Iconic biases across modalities

Jeremy Kuhn, Carlo Geraci, Philippe Schlenker, Brent Strickland

► **To cite this version:**

Jeremy Kuhn, Carlo Geraci, Philippe Schlenker, Brent Strickland. Boundaries in space and time: Iconic biases across modalities. *Cognition*, 2021, 210, pp.104596. 10.1016/j.cognition.2021.104596 . hal-03408801

HAL Id: hal-03408801

<https://hal.science/hal-03408801>

Submitted on 16 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Boundaries in space and time: Iconic biases across modalities

Jeremy Kuhn¹, Carlo Geraci¹, Philippe Schlenker^{1,2}, Brent Strickland^{1,3}

1. Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research

University

2. New York University

3. School of Collective Intelligence, UM6P, Ben Guerir, Morocco

Running Head : Boundaries in space and time

Address for correspondence : Jeremy Kuhn or Brent Strickland
Institut Jean Nicod
Pavillon Jardin
Ecole Normale Supérieure
29, d'Ulm
75005 Paris
France

Email address : jeremy.d.kuhn@gmail.com or brent.strickland@ens.fr

Word count : 10853 (abstract + main text)

Abstract (247 words)

The idea that the form of a word reflects information about its meaning has its roots in Platonic philosophy, and has been experimentally investigated for concrete, sensory-based properties since the early 20th century. Here, we provide evidence for an abstract property of ‘boundedness’ that introduces a systematic, iconic bias on the phonological expectations of a novel lexicon. We show that this abstract property is general across events and objects. In Experiment 1, we show that subjects are systematically more likely to associate sign language signs that end with a gestural boundary with telic verbs (denoting events with temporal boundaries, e.g., *die*, *arrive*) and with count nouns (denoting objects with spatial boundaries, e.g., *ball*, *coin*). In Experiments 2-3, we show that this iconic mapping acts on conceptual representations, not on grammatical features. Specifically, the mapping does not carry over to psychological nouns (e.g. people are not more likely to associate a gestural boundary with *idea* than with *knowledge*). Although these psychological nouns are still syntactically encoded as either count or mass, they do not denote objects that are conceived of as having spatial boundaries. The mapping bias thus breaks down. Experiments 4-5 replicate these findings with a new set of stimuli. Finally, in Experiments 6-11, we explore possible extensions to a similar bias for spoken language stimuli, with mixed results. Generally, the results here suggest that ‘boundedness’ of words’ referents (in space or time) has a powerful effect on intuitions regarding the form that the words should take.

Keywords:

language universals; cognitive biases; event and object boundaries; sign language; telicity; count/mass

Since at least Saussure (1916), linguists and psychologists have acknowledged that the correspondence between the form of a word and its meaning is in part arbitrary. However, work on ‘sound symbolism’ or ‘motivated mappings’ has shown that certain form-meaning pairings are cognitively more natural than others. Investigation of these mappings has yielded insight into the cognitively natural classes with which humans partition semantic space, as well as the evolutionary process by which languages and lexicons develop (Dingemanse et al., 2015).

In the present work, we investigate motivated mappings that involve the abstract notion of *boundariness*. Spatial and temporal boundaries have been a central operating concept in psychological work on object cognition and event cognition (Spelke, 1990; Sharon & Wynn, 2005; Zacks & Swallow, 2007). In language, boundariness and related logical concepts have been central to the characterization of the mass/count and telic/atelic oppositions for nouns and verbs respectively. Intriguingly, boundariness has also been shown to be involved in a motivated mapping in sign language: telic verbs are associated with gestural boundaries (Wilbur, 2003; Strickland et al., 2015). Here, we investigate the origin of this mapping bias, looking at the degree to which it may be abstract and domain-general, and the degree to which it may operate on purely grammatical as opposed to conceptual representations.

Iconicity as a window into cognitive biases

The idea that form-to-meaning (and meaning-to-form) mappings may be in part non-arbitrary (or, ‘motivated’) has roots in Platonic philosophy, and has been experimentally investigated since the start of the 20th century (Köhler, 1929; Davis, 1961; Bremner et al., 2013). For example, high front vowels (e.g. /i/) have been shown to be associated with smallness or closeness (i.e., placement at a small distance) and low back vowels (e.g. /a/) with largeness and

remoteness (Sapir, 1929; Thompson and Estes, 2011). This cognitive tendency is stochastically reflected in the languages of the world in the phonological form of the words ‘small’ and ‘large’ and the words ‘here’ and ‘there’—e.g. French *ici*, ‘here,’ vs. *là*, ‘there’ (Ultan, 1978). Blasi et al. (2016) established similar stochastic tendencies (e.g., /i/ is associated with ‘small’, /n/ is associated with ‘nose’) based on the words for 30 concepts in over 6000 languages.

In both spoken and sign language, one of the pressures that may lead to non-arbitrary mappings is that of *iconicity*: that is, resemblance between properties of a linguistic form and properties of its meaning (Westermann, 1927; Dingemanse, 2012; Cuxac, 2001; Liddell, 2003). To date, descriptive and experimental studies on iconicity have typically focused on mappings between phonological space and qualitative properties that are represented in perception (e.g., size and shape; sound and movement), reflecting the fact that spoken languages often use iconic language to reflect visual and auditory properties (Dingemanse, 2012). Extending this, some recent work has examined the iconic properties of more abstract, logical semantic categories that are linguistically relevant elsewhere in language. Most notably, the phonetic properties of verbs, including length and reduplication, have shown to have interpretive effects relating to verbal aspect, such as duration and iterativity (Dingemanse, 2015; Kuhn & Aristodemo, 2017).

In the present work, we focus on the abstract concept of boundariness. On the psychological side, boundedness is a fundamental property in non-verbal representation relevant to both the spatial domain (objects) and the temporal domain (events). On the linguistic side, syntactic and semantic work has shown strong parallels between count/mass nouns and telic/atelic verbs, intuitively connected to these non-linguistic representations. Boundedness thus provides a highly interesting case for testing mapping biases between form and abstract meaning.

Boundarihood is integral for non-linguistic representations of both space and time

Consider first the spatial domain. Intuitively, *objects* are integral wholes, with stable boundaries and non-homogenous internal structure; *substances* have flexible boundaries and homogenous internal structure. Spelke (1990) argues that infants naturally segment the (spatial) world into discrete objects using boundarihood (where parts of the same object share a single bounded contour). The ability to identify boundaries has been shown to be important for a variety of cognitive tasks, such as infants' ability to track quantities (Cherries et al., 2008). Other work has established a cognitive distinction between objects, which have clear boundaries, and substances, which do not. For example, vanMarle and Wynn (2011) showed that infants were better at tracking quantities of food in the form of bounded objects than unbounded substances. VanMarle and Scholl (2003) showed that this fundamental opposition continues through to adulthood: the adult visual system also has more trouble tracking unbounded substances than bounded objects.

Next consider the temporal domain. Just as a continuous visual display is naturally segmented into discrete objects bounded in space, humans naturally segment continuous activity into discrete *events* bounded in time (Zacks & Swallow, 2007). Like for object perception, event segmentation appears early in development; infants as young as 6 months naturally segment continuous action into temporally bounded events (Hespos et al., 2009; Sharon & Wynn, 1998). As in the spatial domain (where disruption of spatial boundaries impairs object tracking), disrupting temporal boundaries can impede performance on cognitive tasks such as memory encoding (Sonne et al., 2016; Swallow et al., 2009). Finally, in a category identification task, Ji & Papafragou (2020) establish that a cognitive distinction exists between bounded events, which have internal structure and a clear end point (such as folding up a handkerchief), and unbounded events, which have homogenous internal structure and an arbitrary end point (such as waving a

handkerchief). Across a variety of tasks throughout development, bounarihood is thus relevant to the cognitive representation of objects (bounded in space) and events (bounded in time).

The relationship between conceptual boundarihood and linguistic structure

In adults with developed language, there is reason to suspect that the abstract notion of boundarihood exerts a bias on linguistic structure. For nouns, a large body of work has explored the relation between the object/substance opposition and the *count/mass* distinction that appears in many languages. In languages like English, this distinction appears in the way that nouns are quantified: count nouns can be pluralized, and may be quantified with words like *many*; mass nouns cannot be pluralized, but can be quantified in the singular form with the word *much*.

- | | |
|-------------------|----------------|
| (1) too many cats | *too much cat |
| (2) too much mud | *too many muds |

There is a clear and attractive intuition that the count/mass distinction aligns with the object/substance opposition: canonical count nouns refer to objects while canonical mass nouns refer to substances. Indeed, this statistical correlation has been confirmed for a set of 312 nouns commonly found in young children's vocabularies (Samuelson & Smith, 1999). On the other hand, linguistic data within and across languages clearly shows that there cannot be a one-to-one mapping between linguistic and conceptual representations: some concepts may appear in either mass or count syntax (e.g., *too much string* vs. *too many strings*), while other concepts vary across languages (e.g. English *furniture* vs. French *meubles*; see Kulkarni et al., 2013 for a large-scale investigation of cross-linguistic variation). Thus, in many modern linguistic theories of count/mass, the linguistic distinction is not dictated by the conceptual representation but instead

arises via some interaction of the lexical semantics, the syntactic frame, and pragmatic factors, (Bale & Barner 2009; Chierchia 2010; Rothstein, 2010; Srinivasan & Barner, 2020).

The object/substance opposition has nevertheless been shown to exert a *bias* on count/mass categorization in natural language. Brown (1957) showed that 3-5 year old children extend the meaning of a novel word differently depending on whether it is presented in a count or mass syntax: children preferred associating count nouns to unknown objects and mass nouns to unknown substances. Similar word extension studies have repeatedly replicated this finding in paradigms that tease apart the influence of grammatical cues and conceptual categories in children and adults (Gordon, 1985; Soja et al., 1991; Imai & Gentner, 1997; Prasada et al., 2002; *i.a.*). While Gordon (1985) showed that word extension by 3-5 year old children was dominated by grammatical cues (in cases of grammar-concept mismatch), Soja, Carey & Spelke (1991) showed that two-year-old children who have not yet acquired the syntax of count/mass nevertheless extend novel word meanings based on whether something is an object or a substance. Related results have been found even for speakers of languages without a grammatical count/mass distinction (Imai & Gentner, 1997; Inagaki & Barner, 2009). Together, these results suggest that, while languages may vary with respect to lexical semantics and the meaning of specific syntactic constructions, the conceptual representations and categories that ground these distinctions are shared across languages (Srinivasan & Barner, 2020).

Turning to verbs, temporal boundedness is associated with *telicity* (Vendler, 1957; Dowty, 1979; Tenny, 1994). Telicity refers to a temporal property of verb meanings: telic verbs (e.g. *decide*) refer to events bound in time by a natural culmination point; atelic verbs (e.g. *ponder*) refer to events with no such natural culmination. This property is grammatically encoded in natural language. In English, for example, telicity is reflected in the preposition that is chosen

to express the duration of an event: telic predicates take *in*; atelic predicates take *for*.

(3) Eva came to a decision **in** 10 minutes.

(4) Eva pondered the question **for** 10 minutes.

Work on formal semantics has argued for a logical parallel between the semantic representation of telic verbs and count nouns on the one hand and atelic verbs and mass nouns on the other (Bach, 1986; Jackendoff, 1991; Krifka, 1992). Logically, this can be seen in part-whole structure in the two domains: a pondering event can be divided temporally into small subevents, each of which is also a pondering event. However, a deciding event cannot be temporally divided into subevents of deciding, since only the final subevent includes the critical moment of decision (Bennett & Partee, 1972). Analogously, substances like rice can be divided spatially into small subparts, each of which can also be called 'rice', but if a cat is divided spatially into small subparts, these subparts are no longer cats (Cheng, 1973).

As in the nominal domain, languages vary with respect to how verb phrases encode event structure. For bounded events, for example, a temporal endpoint may either be encoded by the verb (*enter the stadium on horseback*) or by a prepositional phrase (*ride into the stadium*). On the other hand, although languages differ with respect to which grammatical strategies are available, non-linguistic tasks have shown that conceptual representations of events remain constant across speakers of different languages (Papafragou et al., 2002; Gennari et al., 2002; Papafragou et al., 2008; Papafragou, 2015). Like for nouns, universally available cognitive representations thus provide a framework on which linguistic structure can be built (see also Strickland, 2016).

A motivated mapping relating to boundarihood

As it turns out, the property of telicity has been shown to be subject to a motivated mapping in sign languages: Wilbur (2003; 2008) observed that telic verbs in American Sign Language (ASL) are systematically signed with a ‘gestural boundary.’ Roughly speaking, telic verbs stop abruptly, often with contact; atelic verbs have a continuous, extendable movement. The precise operationalization of these phonetic cues varies somewhat in the literature. Wilbur (2003) defines ‘end-marking’ using prosodic features of Brentari (1998)’s phonological model: end-marked verbs are those that have a change in aperture, orientation, setting, and/or location. Malaia & Wilbur (2012) define end-marking using fine-grained phonetic features: end-marked verbs have a high peak velocity and rapid deceleration following peak velocity.

Although one can find counter-examples to this generalization (Davidson et al. 2019), a body of literature has established the robustness of the correlation as a strong stochastic effect across several sign languages (Schalber, 2006; Milković, 2011; Malaia & Wilbur, 2012). Malaia & Wilbur (2012) established a correlation between telicity and end-marking for 40 verbs in ASL and 30 verb pairs (60 verbs total) in Croatian Sign Language; Strickland et al. (2015) established the same correlation based on 18 verbs in each of Italian Sign Language, Sign Language of the Netherlands, and Turkish Sign Language. This bias also seems to affect derivational processes. Wilbur (2008), building on Klima and Bellugi (1979), showed that modifications to the phonetic form of an ASL verb may induce aspectual modifications. Wilbur, Malaia, & Shay (2012) showed that similar phonetic manipulations can modify the meaning of adjectives.

Strickland et al. (2015) provided evidence that this mapping bias may be universally accessible by showing that (hearing) subjects who do not know any sign language can nevertheless infer the telicity of a signed verb from its phonetic form. Specifically, subjects were presented with a sign language verb, and were asked to guess which of two English verbs—one

telic and one atelic—corresponded to the meaning of the sign. Even when neither translation was the actual translation of the sign, subjects systematically assigned telic meanings to telic signs more frequently than to atelic signs. Strickland et al. (2015) thus concluded that this correspondence arises from a universally accessible ability to map visual form to event structure.

The origins of motivated mappings relating to boundarihood

Where does this motivated mapping come from? A first possible answer to this question is that the mapping is specific to *gestural boundaries* and *events*. Such a hypothesis seems plausible in light of recent work on the role of the visual system in non-linguistic tasks of event segmentation. Specifically, Zacks et al. (2007) argue that natural points of event segmentation arise via predictive perceptual processing. For visual stimuli, event boundaries roughly correspond to abrupt changes in visually perceptible movement (Zacks et al., 2009). Building on this work, Malaia (2014) argues that sign languages, by virtue of their visual modality, use the same kinematic cues to encode the end-state of telic verbs. These results are consistent with a hypothesis on which the universally-accessible mapping tracks a cognitive function dedicated to visual event perception. This narrow view consequently predicts that the mapping bias should be specific to the visual modality and to temporal boundaries.

In light of the abstract concept of boundarihood discussed above, a second, broader hypothesis acknowledges that motivated mappings between gestural boundaries and events exist, but attributes these to a more general case of representational iconicity. This mapping can be stated in schematic terms rather than in terms of directly perceptible properties: a bounded form is mapped to a bounded meaning and an unbounded form is mapped to an unbounded meaning. This general iconic view consequently predicts that the motivating mapping seen for sign

language verbs should extend to other domains with analogous structure, as suggested by Wilbur, Malaia, & Shay (2012). Regarding meaning, the effect may not be specific to event meanings: a semantic representation may contain a boundary in time or a boundary in space. Regarding form, the effect may not be specific to the visual modality: a phonetic form may contain a boundary in movement or a boundary in acoustic structure. Since existing work on motivated mappings has not investigated the role of boundariness for nouns or for spoken language, both of these predictions are as of yet untested.

In this work, we provide evidence in favor of the latter hypothesis, by establishing that the mapping is general across events and objects. In Experiment 1, we establish that the motivated mapping is not restricted to verbs and events; after replicating Strickland et al.'s results in the verbal domain, we show that an analogous mapping is present in the nominal domain. Inspired by the linguistic parallels between verbs and nouns, we show that a mapping preference holds for gestural stops to count nouns that denote objects bound in space (e.g. *coin*), versus mass nouns that denote substances with no such boundary (e.g. *water*). In Experiments 2 and 3, we show that this iconic mapping is acting on conceptual representations, not on grammatical features. In particular, the mapping does not carry over to psychological nouns (e.g. *idea* vs. *knowledge*), despite the fact that these nouns are still syntactically encoded as either count or mass. We attribute this difference to a representational distinction between physical count nouns and psychological count nouns: the former are bound in space while the latter are not. Experiments 4-5 replicate these findings with a new set of stimuli, controlling for potential confounds. Finally, in Experiments 6-11, we explore possible extensions to a similar bias beyond sign language and the gestural domain. We provide evidence for a similar effect in spoken language with nominal meanings, although interpretation of this finding is left open.

Experiment 1

The work described above establishes that telic meanings are associated with sign language signs that have a gestural boundary. In Experiment 1, we tested the generality of this motivated mapping with respect to meaning. In particular, we tested whether the effect seen for events in the verbal domain extends similarly to the nominal domain. As discussed above, work on formal semantics has drawn deep parallels between the telic/atelic distinction in the verbal domain and the count/mass distinction in the nominal domain: the abstract notion of boundariness that distinguishes *decide* from *ponder* is parallel to the one that distinguishes *thing* from *stuff*. Motivated by these formal parallels, here we tested whether the phonetic properties that allow subjects to infer that a signed verb is telic or atelic are also spontaneously employed by subjects to infer whether a noun is count or mass.

In order to test the presence of a motivated mapping, we adopted the experimental method of Strickland et al. (2015) in which participants were asked to match a sign language verb to one of two English translations, differing with respect to telicity. Stimuli were videos of verbs in Italian Sign Language (LIS). LIS signs were identical to those of Strickland et al. (2015), in which verbs were classified as telic or atelic by a Deaf native signer, and a separate rating study was carried out to confirm the presence or absence of a "gestural boundary." In replicating the experimental design of Strickland et al. (2015) as closely as possible, we consequently also adopt Strickland et al.'s operationalization of end-marking as a holistic property corresponding to whatever phonetic features are characteristic of telic verbs, a property tracked in non-signers by perception of a "gestural boundary." (Experiments 4 and 5 control potential confounds with a different operationalization of end-marking.) In Experiment 1a, meaning choices were verbs denoting physical actions (e.g. *slip* vs. *limp*). In Experiment 1b, meaning choices were nouns

denoting physical objects (e.g. *coin* vs. *rain*). To verify that the results reported in Strickland et al. (2015) were not a result of the specific meaning choices used in that study, Experiment 1a used an entirely new set of verb meanings. More generally, there was no overlap between the 18 meaning choices that appear in Strickland et al. (2015) and the 79 meaning choices that appear in the experiments described in the present paper.

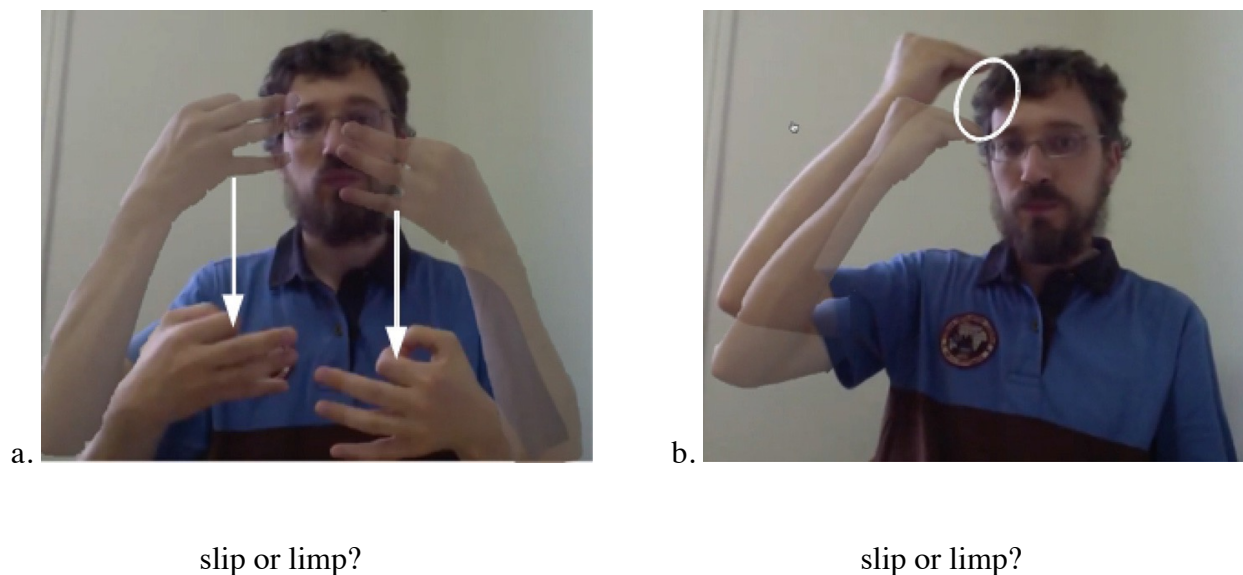
Methods

Participants. An identical experimental method was run in Strickland et al. (2015), Experiment 3 with 22 participants at 0.96 power. Here we decided to multiply the number of participants by five, leaving the studies intentionally over-powered, since we were unsure of how effect sizes might change according to domain and of what effect size to expect for any interactions. We thus requested 100 participants through Amazon's Mechanical Turk; when more than 100 subjects responded, we kept all available data. Experiment 1a analyzed responses from 97 subjects. An additional eight subjects completed the experiment but were excluded from analysis because they failed an attention check at the beginning of the study or because they later listed a native language other than English. Experiment 1b included responses from 99 subjects; an additional eight were excluded from analysis for attention check failure or native language. Gender and age information for all experiments is included in the supplemental materials. The studies reported were approved by an IRB panel at the CERES in Paris, France, and all participants provided informed consent.

Materials and Procedure. Each participant was asked to guess the meaning of 18 signs from a pair of meanings. Signs were videos of verbs in LIS. Of these 18 forms, nine were telic verbs (displaying a gestural stop) and nine were atelic (without a gestural stop). Subjects were

instructed to view each video as many times as needed. For Experiment 1 and all subsequent experiments, all stimuli and meaning choices are presented in Appendix A.

Stimuli in both experiments were presented one at a time and in a randomized order, with stop stimuli mixed with non-stop stimuli. Meaning choices consisted of nine pairs of English verbs, roughly matched for word length. In Experiment 1a, meaning choices were verbs denoting physical events; each pair consisted of one telic and one atelic verb. In Experiment 1b, meaning choices were nouns denoting physical objects; each pair consisted of one count and one mass noun. Neither of the meaning choices corresponded to the actual meaning of the verb (in the case of signs). Each of these nine pairs of meanings was presented once for a stop-stimulus and once for a non-stop-stimulus. The association of stimulus and meaning-pair was randomized to generate four lists of test items for each experiment. For two lists, four of the nine meaning-pairs presented the telic/count meaning as the first answer. The order in which the meanings appeared was reversed for the other two lists. Each subject saw only one list total. Example stimuli are shown in Figure 1.



coin or rain?

coin or rain?

Figure 1: Example stimuli for Exp. 1a (verb choices) and Exp. 1b (noun choices). Figure 1a is a LIS sign with a gestural stop ('decide'); Figure 1b is a LIS sign without a gestural stop ('think').

Results

In Experiment 1 and subsequent experiments, we used logit mixed-effects models to model the likelihood of the response type (Jaeger, 2008)—here, whether the telic meaning or the atelic meaning was selected. Participants, stimuli, and meaning choices were included as random intercepts (random slopes often did not converge). For each subexperiment, gestural boundariness (i.e. the presence or absence of a boundary in a stimulus) was added as a predictor variable to see whether the model was improved. When experimental results were combined to test interactions, predictor variables (e.g., gestural boundariness; part of speech of meaning choices) were added incrementally to see whether the model was improved. Model fit was assessed using chi-square tests on the log-likelihood values of competing models. In general, we report only results for theoretically motivated effects and interactions—namely, gestural boundariness and its interaction with other predictor variables. (Appendix B provides full summaries of maximal models for Experiments 1-7.)

Both Experiment 1a and Experiment 1b showed a main effect with respect to presence of a boundary. In both the verbal and nominal domain, participants provided more telic/count responses for stop-stimuli than for non-stop-stimuli (Exp 1a: 63.5% vs. 35.5%, $\chi^2(1) = 9.37$, $p = 0.002$; Exp 1b: 58.4% vs. 26.9%, $\chi^2(1) = 9.89$, $p = 0.002$). Combining responses from Experiments 1a and 1b, no interaction was found between boundariness and part of speech ($\chi^2(1) = 1.37$, $p = 0.24$).

Discussion

Experiment 1a establishes the robustness of the motivated mapping observed in Strickland et al. (2015) by replicating these results with 18 novel verb meanings. Experiment 1b establishes that the meanings involved in this motivated mapping are not isolated to verbs and events. If the effect were driven by properties specific to events, then no effect would be predicted for nouns meanings. The fact that analogous results hold for count and mass nouns thus shows that the factor at play is not a property specific to events, but rather a domain-general representation of boundariness.

Experiment 2

As observed in the introduction, the mapping between the object/substance opposition and the count/mass distinction is not one-to-one. Of particular note, many concepts show cross-linguistic variation in their syntactic encoding; for example, *spinach* is mass in English but *épinard(s)* is count in French. Inagaki & Barner (2009) provide evidence that concepts that vary in their grammatical encoding across languages (such as *spinach*, *hair*, *pasta*, *toast*) are precisely those that are unclear in their conceptual categorization as objects or substances. In Experiment 1, we deliberately tested canonical, concrete verbs and nouns (e.g., object-denoting count nouns vs. substance-denoting mass nouns), in order to cleanly isolate the extension from the verbal domain to the nominal domain. A consequence of this decision is that the conceptual and linguistic variables were confounded: is the iconic mapping acting on grammatical features or on conceptual representations?

This question can be addressed by focusing on words where the correspondence breaks down: cases in which a word belongs to a clear grammatical category but does not have a clear

conceptual categorization. We hypothesized that this may systematically be the case for certain types of *abstract* words. In particular, psychological verbs and nouns may be grammatically categorized in a given language with respect to telicity or count/mass (e.g. *many ideas* vs. **many knowledges*), even if the conceptual boundaries that are involved for psychological events and objects are not directly perceivable.

Even more specifically, there is an important intuitive difference between the psychological entities referred to by nouns and the psychological events referred to by verbs. For nouns that refer to physical objects, boundedness can often be defined in terms of spatial properties of the referent like contiguity, size, and topology (Sutton & Filip, 2016). In contrast, there is no clear sense in which ideas (but not knowledge) are spatially bounded. Psychological count nouns (e.g. *idea*) are thus unlikely to be conceptually represented with a boundary in space. On the other hand, it seems entirely possible that psychological telic verbs (e.g. *decide*), like their concrete counterparts (e.g. *break*), are represented as having a boundary in time.

Experiment 2 thus sought to dissociate the role of grammatical categories from the role of conceptual representations in the iconic mapping observed in Experiment 1. We define two refinements of the iconicity hypothesis. On the first sub-hypothesis, the iconic bias is mediated by the linguistic representation, mapping between grammatical categories and form. On this hypothesis, the effects seen above should be identical in the physical and psychological domains, directly reflecting grammatical categorization. On the second sub-hypothesis, the iconic mapping bias involves non-linguistic conceptual representations, independent of grammatical facts. On this sub-hypothesis, the iconic bias should hold to the extent that boundariness is relevant in the conceptual representations. Based on the intuitions above, we would thus predict to see a complex interaction, in which the iconic mapping disappears for psychological nouns. These

predictions were addressed in Experiment 2.

In Experiment 2, stimuli were videos of LIS verbs. In Experiment 3a meaning choices were pairs of psychological verbs (e.g. *choose* vs. *dream*). In Experiment 3b, meaning choices were pairs of psychological nouns (e.g. *concept* vs. *contempt*).

Methods

Participant recruitment and exclusion criteria were identical that of Experiment 1. Experiment 2a included responses from 98 subjects; an additional three were removed from analysis. Experiment 2b included responses from 101 subjects; an additional five were removed from analysis. The procedure and materials were identical to those of Experiment 1, except in the following way: meaning choices in Experiment 2 were psychological verbs and nouns.

Results

Analysis methodology was identical to that of Experiment 1. Experiment 2a showed a main effect with respect to presence of a boundary. Participants provided more count responses for stop-stimuli than for non-stop-stimuli (64.3% vs. 39.3%, $\chi^2(1) = 9.25$, $p = 0.002$). Experiment 2b produced results that were different in character from all preceding experiments. Experiment 2b showed a weak main effect with respect to the presence of a boundary, but this effect was in the opposite direction from all other experiments: participants provided fewer telic responses for stop-stimuli than for non-stop-stimuli (52.3% vs. 61.5%, $\chi^2(1) = 2.8047$, $p = 0.094$). Combining responses from Experiments 2a and 2b, an interaction was found between boundariness and part of speech ($\chi^2(1) = 112.02$, $p < 0.001$).

Boundary x Part of speech x Conceptual domain

The fact that the presence of a boundary generates an effect for all verbs but only physical nouns was confirmed out by a significant three-way interaction between boundedness (boundary vs. no boundary), part of speech (noun vs. verb) and conceptual domain (physical vs. psychological). Combining responses from Experiments 1 and 2, we found a significant three-way interaction ($\chi^2(3) = 119.85, p < 0.001$). Psychological nouns differ from physical nouns and all verbs in being the only condition that is not sensitive to the presence of a boundary in the stimulus. Summaries of the maximal models for these and later experiments can be found in Appendix B.

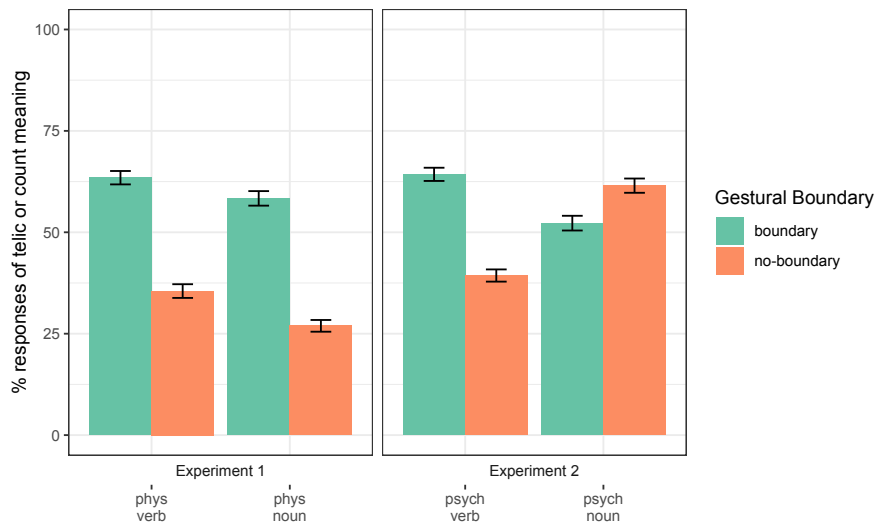


Figure 2. Results for Experiments 1 and 2. Error bars indicate standard error by participant.

"Boundary" stimuli (green bars) indicate LIS signs with a gestural boundary.

Discussion

Experiment 1 established an iconic mapping for verbs and nouns denoting physical events and entities. While Experiment 2a showed an identical effect for verbs referring to psychological events, Experiment 2b failed to demonstrate a mapping for nouns referring to

abstract psychological entities. This pattern of results reflects the intuition that psychological nouns lack a boundariness distinction in the conceptual representation, despite being clearly categorized as count or mass. In contrast, the data are not expected on a theory in which the effect is mediated by grammatical features, which would predict no difference between physical and psychological nouns. These results thus support the hypothesis that the iconic mapping acts on conceptual representations and not on grammatical features.¹

Experiment 3

The predictions in Experiment 2 were based on the intuition that psychological nouns are not conceptually represented with a clear boundary in either time or space, but that psychological verbs may still be represented with a temporal boundary. To date, however, there is no existing empirical work that directly compares the conceptual representation of boundariness across events and objects. To establish the robustness of the intuitions above, we thus performed four experiments on the 72 words tested in Experiments 1-2 plus seven additional psychological nouns tested in Experiment 5. We tested conceptual representations of word meanings by asking separate pools of subjects to evaluate properties of the events or entities described by a given word. Experiment 3a tested verb meanings for boundariness; Experiment 3b did the same with noun meanings. Experiment 3c tested verb meanings for divisiveness; Experiment 3d did the same with noun meanings.

Methods

Participants. In each of the experiments, we requested 20 participants through MTurk,

¹ This notably diverges from the linguistic claims of Wilbur (2003), who proposes that, for native sign language users, end-marking is precisely a grammatical feature (called *res*). But importantly, these claims are made in two different domains: by looking at non-signers, we are investigating the cognitive underpinnings of this mapping; we make no claims about the way these phonetic features may be grammaticalized within a developed linguistic system.

and removed those who failed an attention check. Participant numbers were thus as follows.

Experiment 3a: 18 subjects; an additional two were excluded. Experiment 3b: 16 subjects; an additional four were excluded. Experiment 3c: 18 subjects; an additional two were excluded.

Experiment 3d: 16 subjects; an additional five were excluded.

Materials and Procedure. The list of word meanings tested in Experiment 3 consisted of exactly the 79 meaning choices used in the other experiments in the present paper, provided in Appendix A.

In Experiments 3a and 3b, conceptual objects and events were operationalized in terms of boundariness, following insights from the psychological literature. Each subject was asked to consider an entity or event that could be described with a given word, and then to rate on a scale from 1 to 7 the degree to which it has a "clear and stable boundary." In Experiments 3c and 3d, conceptual objects and events were operationalized in terms of divisiveness, following insights from the formal semantics literature. Each subject was asked to consider an entity or event that could be described with a given word, and then to rate on a scale from 1 to 7 the degree to which the same word could apply to *parts* of what was described. An example item is provided in (5). The two operationalizations were expected to behave similarly and to provide a measure of the boundedness for each entity or event concept.

- (5) Consider something that you could describe with the following word:

coin

Exp 3b: Does what is described have a **clear and stable boundary**?

Exp 3d: Can the same word apply to **parts** of what is described?

Results

For the four experiments, we performed an ordinal regression to model the likelihood of the judgment assigned on the 7-point scale. Participants and items were included as random intercepts; predictor variables (count/telic vs. mass/atelic; noun vs. verb; physical vs. psychological) were added incrementally to see whether the model was improved. Model fit was assessed using chi-square tests on the log-likelihood values of competing models.

Confirming our informal intuitions, we found a significant effect of grammatical category (count/telic vs. mass/atelic) on the conceptual measure for all groups except for psychological nouns. The same pattern of results was found for both operationalizations of the conceptual property: as boundariness or as divisiveness. (Exp 3a: Physical verb boundariness: $\chi^2(1) = 16.12$, $p < 0.001$; Psychological verb boundariness: $\chi^2(1) = 6.03$, $p = 0.014$. Exp 3b: Physical noun boundariness: $\chi^2(1) = 40.44$, $p < 0.001$; Psychological noun boundariness: $\chi^2(1) = 1.41$, $p = 0.24$. Exp 3c: Physical verb divisiveness: $\chi^2(1) = 23.89$, $p < 0.001$; Psychological verb divisiveness: $\chi^2(1) = 11.90$, $p = 0.001$. Exp 3d: Physical noun divisiveness: $\chi^2(1) = 21.24$, $p < 0.001$; Psychological noun divisiveness: $\chi^2(1) = 1.57$, $p = 0.21$.) This pattern was confirmed in a significant three-way interaction between category, part of speech, and conceptual domain (boundariness, Exp3ab: $\chi^2(3) = 48.4$, $p < 0.001$; divisiveness, Exp3cd: $\chi^2(3) = 14.49$, $p = 0.002$).

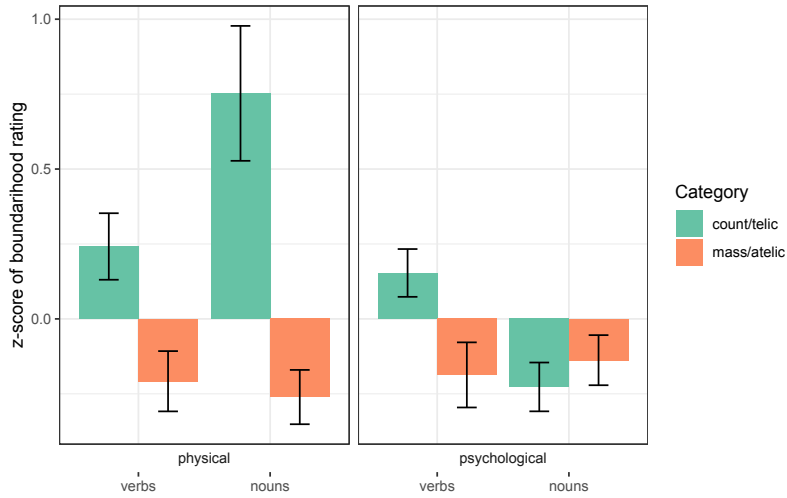


Figure 3. Experiments 3a, b: Z-scores of boundariorhood ratings. Error bars indicate standard error by participant.

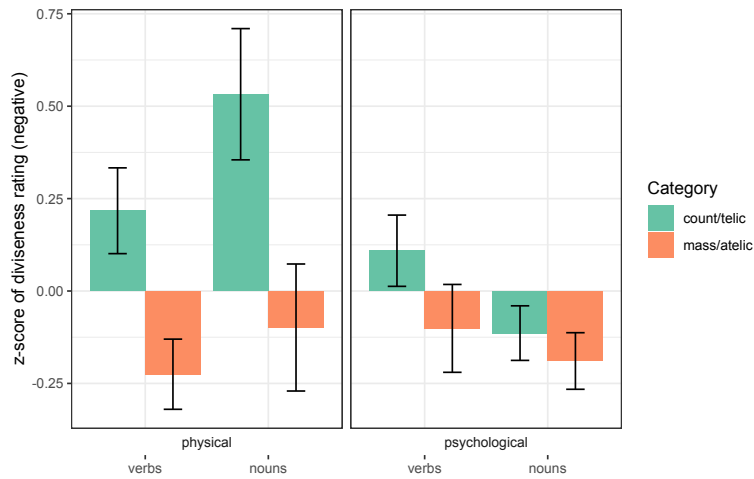


Figure 4. Experiments 3c, d: Z-scores of divisiveness ratings, inverted since boundariorhood and divisiveness are in an inverse relation. Error bars indicate standard error by participant.

Discussion

The telic verbs in our study, whether physical or psychological, denote non-divisible

events with clear temporal boundaries, in contrast with the atelic verbs, which denote divisible events without clear boundaries. Similarly, the physical count nouns in our study denote non-divisible objects with clear boundaries, in contrast with the physical mass nouns, which denote divisible substances without clear boundaries. The psychological nouns in our study are different: no difference was found between the conceptual representation of count nouns and of mass nouns.² Psychological nouns thus provide a systematic natural class that can be used to dissociate grammatical features from conceptual properties.

In Experiments 1-2, if the observed mapping bias were mediated by the linguistic representation, then the mapping bias should be identical in the physical and psychological domains, directly reflecting grammatical categorization. Experiment 2 showed that this was not the case. In contrast, if the mapping bias involved non-linguistic conceptual representations, then the iconic bias should hold to the extent that boundariness is relevant to these representations. Experiment 3 directly measured the extent to which boundariness is a relevant conceptual property for the verbs and nouns in the present study. For two different measures, we saw that boundariness is relevant to verbs (whether physical or psychological) and to physical nouns, but not to psychological nouns. These results directly mirror the interaction observed in Experiments 1-2, thus supporting the hypothesis that the iconic mapping acts on conceptual representations and not on grammatical features.

Experiments 4 and 5

An alternative explanation for the absence of an effect for psychological nouns is that the

² Interestingly, both categories of abstract nouns showed the same general conceptual properties as mass nouns, which supports the position that mass nouns form a more heterogeneous semantic category than count nouns (Srinivasan & Barner, 2020; Kulkarni et al., 2013).

psychological nouns used in Experiment 2b have more ambiguous grammatical categorizations than the physical nouns used in Experiment 1b. As previously mentioned, some nouns can be used in both mass and count frames (e.g. *a lot of string* vs. *many strings*; *a lot of thought* vs. *many thoughts*). Because of their intermediate conceptual representation, it is likely that psychological nouns show more frequent grammatical ambiguity than concrete physical nouns. Although the meaning choices used in Experiment 2b were selected to be as unambiguous as possible with respect to their grammatical categorization, it is possible that the psychological nouns were nevertheless more ambiguous than the physical nouns in Experiment 1b, and that this grammatical ambiguity was responsible for the disappearance of the iconic bias.

Additionally, in previous experiments, sign language stimuli were actual words from LIS. As a consequence, they (a) may have other unremarked but relevant phonetic properties (b) were not finely controlled to form minimal pairs. These facts may introduce noise and unnoticed biases.

In Experiments 4 and 5, we replicated the results from Experiments 1 and 2 while controlling for these possible confounds. For gestural stimuli, minimal pairs of nonce signs were generated by a fluent (non-native) signer, controlling for handshape, position in space, orientation, and duration of the signs across pairs. Here, we operationalized end-marking using Wilbur (2003)'s original characterization in terms of Brentari's (1998) model: the "boundary" stimuli were those that ended with a change of aperture, orientation, or location. In Experiment 4a, meaning choices were physical verbs; in Experiment 4b, meaning choices were physical nouns. In Experiment 5a, meaning choices were psychological verbs; in Experiment 5b, meaning choices were psychological nouns. Noun meaning choices were normed based on grammaticality judgments from an independent pool of subjects, and new meaning choices were chosen for

Experiment 5b (psychological nouns) that were on average *less* ambiguous than the physical nouns in Experiment 4b. In addition to controlling for the confounds described above, the entirely new video stimuli and the new meaning choices for Experiment 5b were intended to increase our confidence in the robustness of the effects and interactions found earlier.

Methods

Participant recruitment was identical that of Experiment 1. In addition to exclusion criteria from Experiment 1, participants in Experiments 4 and 5 were excluded if they reported familiarity with a sign language. Participant numbers were as follows. Experiment 4a: 86 participants; an additional 15 were excluded from analysis. Experiment 4b: 88 participants; an additional 11 were excluded. Experiment 5a: 73 participants; an additional 27 were excluded. Experiment 5b: 72 participants; an additional 27 were excluded.

All noun meaning choices from Experiments 1 and 2 were normed based on grammaticality judgments. To norm noun meanings, subjects were asked to rate the naturalness of a noun in a given grammatical frame on a scale from 1 to 7. A pool of 50 subjects rated the naturalness of nouns in the frame "a lot of X"; a second pool of 46 subjects (50 – 4 excluded) rated the same nouns in the frame "many Xes."³ Subtracting the first rating from the second provided a score of the overall "countiness" for each noun from –6 to +6. For the stimuli used in Experiments 1 and 2, the average overall scores for count and mass nouns in the physical domain were indeed slightly farther apart than the scores in the psychological domain (8.14 vs. 7.01 point difference of "countiness" between the two meaning choices). In order to avoid the confound of

³ Subjects were presented with stimuli such as the following:

(i) **many attentions**
How natural does this sound? (Not natural at all 1 2 3 4 5 6 7 Very natural)

having more ambiguous meaning choices for the psychological nouns, new psychological nouns were chosen that flipped this factor (8.27 vs. 8.99 point difference on a second norming study).

All norming scores are included in supplemental materials.

The procedure and materials were identical to those in Experiment 1 except that sign stimuli were controlled as described above and the new psychological noun meanings were used.

Example stimuli are shown in Figure 4.

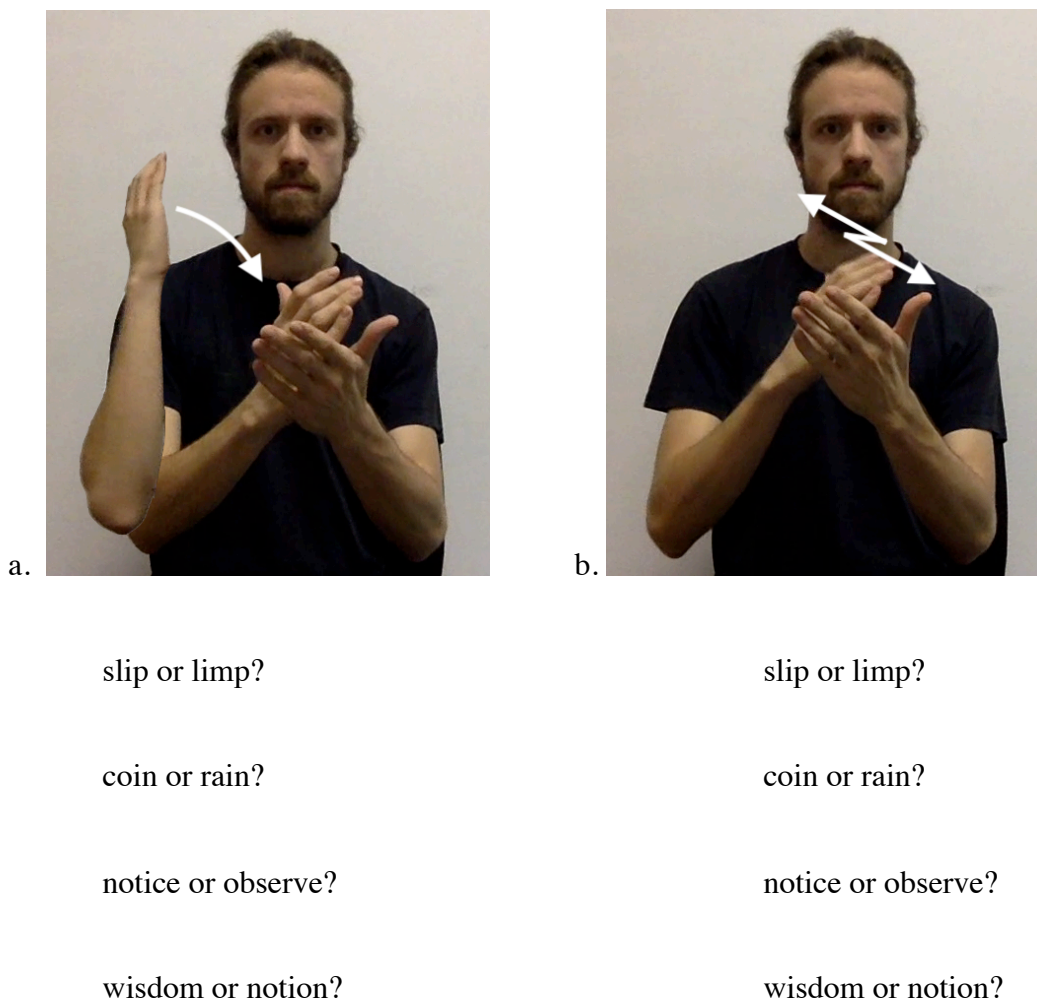


Figure 5: Example stimuli for Exp. 4a (physical verbs) and Exp. 4b (physical nouns), Exp. 5a (psychological verbs) and Exp. 5b (psychological nouns).

Results

Analysis methodology was identical to that of Experiment 1. Experiments 4a and 4b each showed a main effect with respect to presence of a boundary. For sign language nonce words with both physical nouns and physical verbs, subjects provided more telic and count responses for stop-stimuli than for non-stop-stimuli (Exp 4a: 70.5% vs. 39.4%, $\chi^2(1) = 21.01$, $p < 0.001$; Exp 4b: 52.3% vs. 37.8%, $\chi^2(1) = 6.81$, $p = 0.009$). Combining Experiments 4a and 4b, a significant interaction was found between boundedness and part of speech ($\chi^2(1) = 24.79$, $p < 0.001$): nouns showed a smaller effect than verbs. Experiment 5a, with psychological verbs, also showed a main effect with respect to presence of a boundary. Participants provided more telic responses for stop-stimuli than for non-stop-stimuli (60.58% vs. 44.14%, $\chi^2(1) = 8.85$, $p = 0.003$). This effect disappeared in Experiment 5b, with psychological nouns: no effect was found with respect to the presence of a boundary (56.17% vs. 52.93%, $\chi^2(1) = 1.3318$, $p = 0.25$). Combining Experiments 4a and 4b, a significant interaction was found between boundedness and part of speech ($\chi^2(1) = 12.09$, $p < 0.001$): nouns showed a smaller effect than verbs. Combining Experiments 4 and 5, a significant interaction was found between boundedness and conceptual domain ($\chi^2(1) = 24.94$, $p < 0.001$). For both nouns and verbs, there was a reduced effect of the presence of a boundary in the stimulus, and for psychological nouns, the effect disappeared completely.

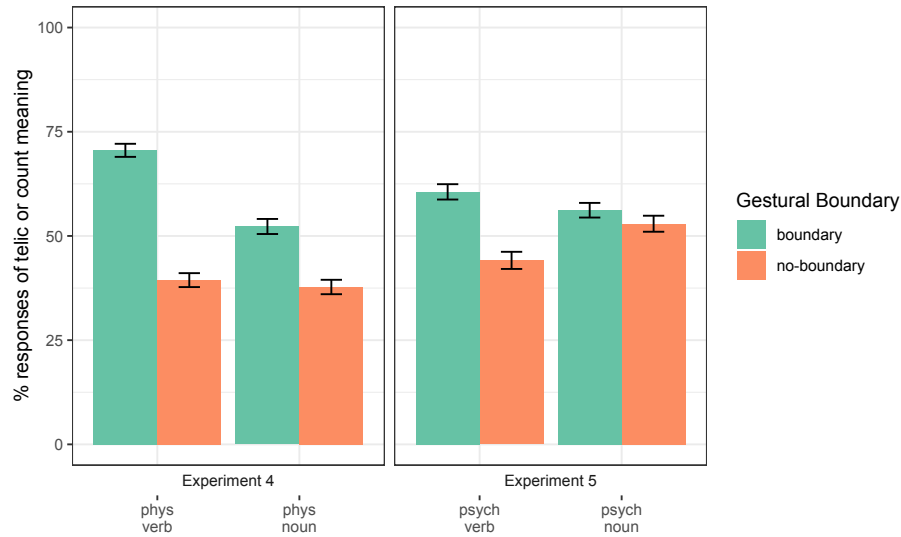


Figure 6. Results for Experiments 4-5. Error bars indicate standard error by participant.

"Boundary" stimuli (green bars) indicate nonce signs with a gestural boundary.

Discussion

In the sign language/gestural modality, Experiments 4 and 5 reproduced results from Experiments 1 and 2 with more tightly controlled stimuli and meaning choices. In the physical domain, the presence of a boundary in the stimuli systematically led subjects to choose more telic or count meanings, but in the psychological domain, this bias disappeared for psychological nouns. In Experiment 5b, with psychological nouns, the bias disappeared in spite of the fact that the grammatical categorization of the psychological nouns was less ambiguous than physical nouns that did show a significant effect in Experiment 4b. These results provide evidence that the iconic mapping bias is sensitive to the properties of the conceptual representations, and not only to the grammatical properties of the words that are used to describe them.

Experiments 6 and 7

In Experiments 1-5, we established that motivated mapping is domain-general with respect to meaning: bounded gestural stimuli were associated with telic verbs denoting events bound in time, and also with count nouns denoting objects bound in space. In the remaining experiments, we examine the extent to which the motivated mapping is domain-general with respect to form, by exploring whether a similar effect can be found for written or audio stimuli. In sign languages, telicity is associated with two phonetic features of the signs: the presence or absence of a ‘gestural boundary’ (more precisely, rapid deceleration of the motion of the hand; see Malaia & Wilbur, 2012), and presence or absence of repetition. In the spoken language domain, no existing work on sound symbolism has investigated the property of boundedness. Nevertheless, the hypothesis of generalized iconicity makes predictions about where this iconic mapping should be found. In particular, plosives are sounds that stop airflow through the mouth; fricatives are sounds that allow continued airflow. We thus constructed written stimuli containing phonological features that correspond to the gestural features in the sign language stimuli. In analogy to the presence or absence of a gestural boundary, written stimuli were chosen that end in a plosive—{p, b, t, d}—or a fricative—{f, v, s, z}. Phonetically, plosives involve complete closure of the air passage and result in an acoustic burst; fricatives involve partial closure and result in an extendable hiss. Repetition of a sign was translated as repetition of the rhyme of a one syllable written word. Pairs of stimuli across the two categories were otherwise identical. Thus, written stimuli contrasted pairs like *bip* vs. *bifif*.

In Experiment 6a, meaning choices were physical verbs (i.e. verbs referring to physical events); in Experiment 6b, meaning choices were physical nouns (i.e. nouns referring to physical entities). In Experiment 7a meaning choices were pairs of psychological verbs; in Experiment 7b, meaning choices were pairs of psychological nouns.

Methods

Participant recruitment and exclusion criteria were identical that of Experiment 1. Participant numbers were as follows. Experiment 6a: 111 subjects; an additional four were excluded. Experiment 6b: 101 subjects; an additional one was excluded. Experiment 7a: 114 subjects; an additional six were excluded. Experiment 7b: 96 subjects; an additional four were excluded. The procedure and materials for Experiments 6 and 7 were identical to those of Experiments 1 and 3, except that phonological forms were written nonce words using English orthography.

Results

Analysis methodology was identical to that of Experiment 1. Both Experiment 6a and Experiment 6b showed a main effect with respect to the presence of a boundary in the written stimuli. For both physical verb meanings and physical noun meanings, participants provided more telic and count responses for stop-stimuli than for non-stop-stimuli (Exp 6a: 56.4% vs. 44.2%, $\chi^2(1) = 6.25$, $p = 0.012$, Exp 6b: 57.8% vs. 44.6%, $\chi^2(1) = 17.27$, $p < 0.001$). Combining responses from Experiments 6a and 6b, no interaction was found between boundariness and part of speech ($\chi^2(1) = 0.048$, $p = 0.83$). Experiment 7a showed a main effect with respect to presence of a boundary for psychological verbs. Participants provided more telic responses for stop-stimuli than for non-stop-stimuli (55.5% vs. 45.9%, $\chi^2(1) = 7.71$, $p = 0.006$). This effect disappeared in Experiment 7b for psychological nouns: participants did not provide more count noun responses for stop than for non-stop stimuli (56.4% vs. 53.0%, $\chi^2(1) = 0.89$, $p = 0.35$). Combining responses from Experiments 8a and 8b, an interaction was found between boundariness and part of speech ($\chi^2(1) = 4.11$, $p = 0.043$). Combining Experiments 6 and 7, a significant interaction was found

between boundedness and conceptual domain ($\chi^2(1) = 6.70, p = 0.010$), with a reduced effect for psychological verbs and nouns.

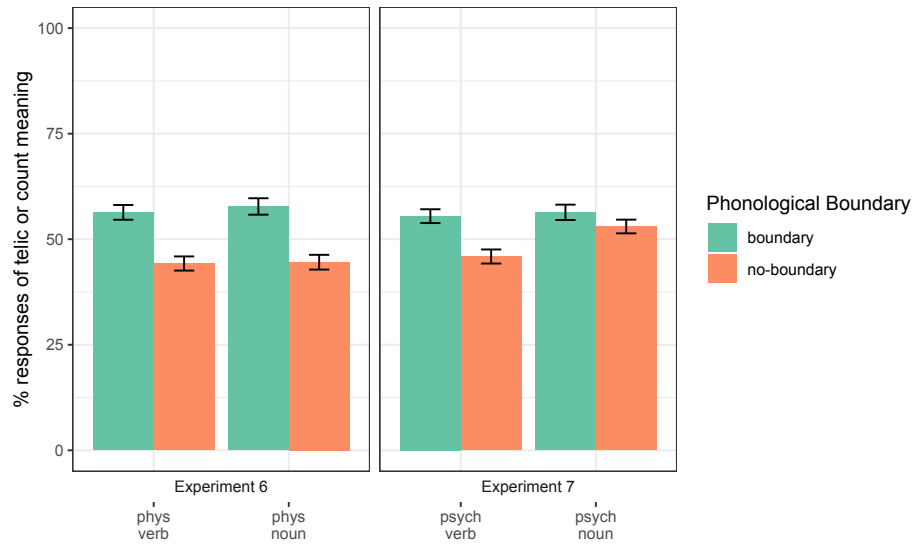


Figure 7. Results for Experiments 6-7. Error bars indicate standard error by participant.

"Boundary" stimuli (green bars) indicate written nonce words ending with a phonological stop.

Discussion

While previous experiments focused on the generality of the *meanings* involved in the motivated mapping system, Experiments 6 and 7 established that the *forms* involved in the motivated mapping observed by Strickland et al. (2015) are not dependent on the gestural modality. Strikingly, exactly the same pattern of results was obtained for written nonce words in Experiments 6 and 7 as was observed for gestural stimuli in all previous experiments: an effect is found in all conditions except for psychological nouns. The fact that analogous results hold for written nonce words with phonological stops suggests that the factor at play is not a property specific to sign language, but rather a domain-general property of the boundedness of the form of

a word.

Experiments 8-11

Experiments 6-7 assumed a correspondence between properties of the form of signs and properties of the form of written nonce words. In particular, we have assumed that gestural stops correspond to acoustic stops (p, t, k) and that repetition of a movement corresponds to repetition of a syllable in a written nonce word. However, these assumptions have introduced several confounds to our design. First, since the nonce words were presented as written stimuli instead of audio stimuli, the form of the written word (i.e., the shape of a *p* vs. an *f*) may introduce additional biases (Cuskley, Simner & Kirby, 2015). Second, the assumed correspondence confounds the three related properties of repetition, number of syllables, and length. In particular, many instances of repetitive or trilled motion in sign language are assumed in the literature to be mono-syllabic (Brentari 1998), just as the multiple flaps of a trilled Spanish /r/ correspond to a single articulatory unit. Signs and nonce words with repetition will also generally have a longer duration; existing work has shown that word length may itself introduce a bias associated with semantic complexity (Lewis & Frank, 2016).

We controlled these confounds by removing each confounding variable in a series of four experiments. In Experiment 8, like in Experiments 6 and 7, length, repetition, and presence of a stop were varied together, but auditory stimuli were used instead of written stimuli, with a different set of items (e.g. /gop/ vs. /pofof/). In Experiment 9, repetition and presence of a stop varied together, but length was controlled by making all stimuli bisyllabic (e.g. /kalop/ vs. /pofof/). In Experiment 10, the contribution of the stop consonant was isolated with monosyllabic stimuli (e.g. /gop/ vs. /gof/). In Experiment 11, the contribution of length was isolated; all stimuli

ended with a stop (e.g. /gop/ vs. /kalop/). In Experiments 8a, 9a, 10a, and 11a, meaning choices were physical verbs; in Experiments 8b, 9b, 10b, and 11b, meaning choices were physical nouns.

Methods

Participant recruitment was identical that of Experiment 1. In addition to exclusion criteria from Experiment 1, participants in Experiments 8-11 were excluded if they failed a second attention check presented auditorially. Participant numbers were as follows. Experiment 8a: 91 subjects; an additional 10 were excluded. Experiment 8b: 82 subjects; an additional 18 were excluded. Experiment 9a: 96 subjects; an additional six were excluded. Experiment 9b: 88 subjects; an additional 10 were excluded. Experiment 10a: 88 subjects; an additional 12 were excluded. Experiment 10b: 89 subjects; an additional 10 were excluded. Experiment 11a: 91 subjects; an additional nine were excluded. Experiment 11b: 85 subjects; an additional 14 were excluded. The procedure and materials were identical to those of Experiment 6 except that audio sound stimuli were controlled as described above.

Results

Analysis methodology was identical to that of Experiment 1. In the verbal domain, Experiments 8a, 9a, 10a and 11a showed no significant effect of either the presence of a phonological stop or of length of the stimulus (Exp 8a: 51.7% vs. 47.9%, $\chi^2(1) = 2.52$, $p = 0.11$; Exp 9a: 54.0% vs. 49.5%, $\chi^2(1) = 0.91$, $p = 0.34$; Exp 10a: 50.1% vs. 53.0%, $\chi^2(1) = 0.63$, $p = 0.43$; Exp 11a: 54.0% vs. 49.5%, $\chi^2(1) = 1.19$, $p = 0.28$). In the nominal domain, Experiments 8b, 9b, and 10b showed a main effect with respect to presence of a phonological stop. Subjects provided more count responses for stop-stimuli than for non-stop-stimuli (Exp 8b: 57.7% vs. 38.9%, $\chi^2(1) = 18.46$, $p < 0.001$; Exp 9b: 59.6% vs. 40.1%, $\chi^2(1) = 22.44$, $p < 0.001$; Exp 10b:

55.9% vs. 43.8%, $\chi^2(1) = 11.63$, $p < 0.001$). Experiment 11b showed no significant effect of the length of the stimulus (52.7% vs. 48.9%, $\chi^2(1) = 0.70$, $p = 0.40$).

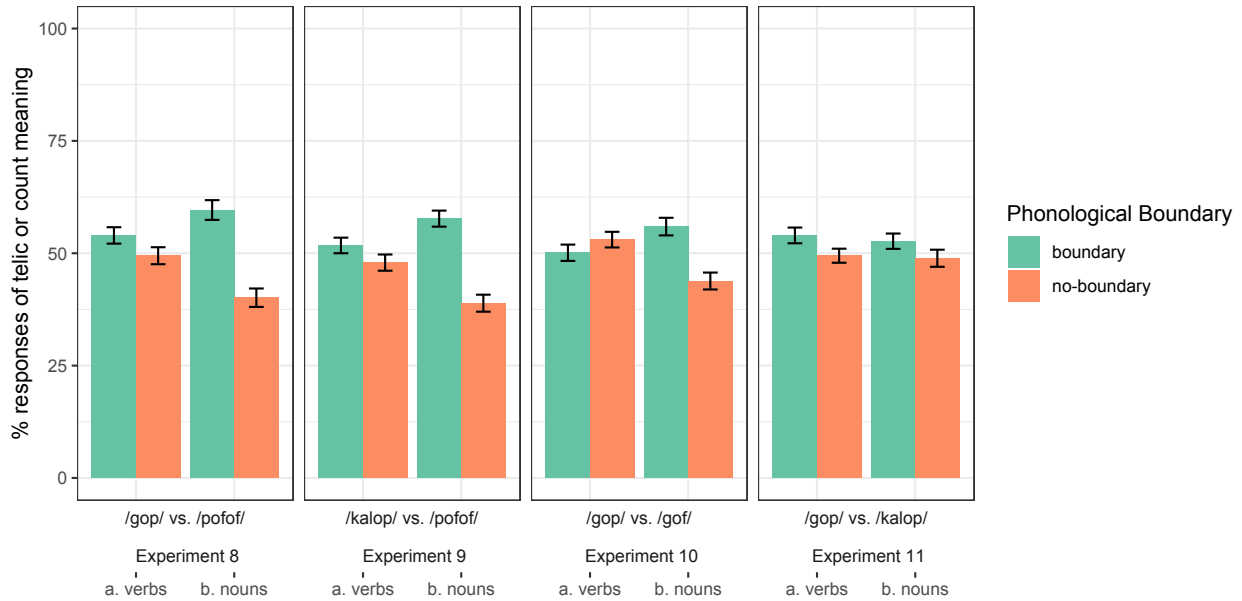


Figure 8. Results for Experiments 8-11. Error bars indicate standard error by participant.

"Boundary" stimuli (green bars) indicate audio nonce words hypothesized to associate with telic or count meanings: stop-final (Exp. 8, 9, 10), with less repetition (Exp. 8, 9), and fewer syllables (Exp. 8, 11).

Discussion

In the nominal domain, the Experiments 8-10 reproduce the results of Experiment 6 (written *bip* vs. *bifif*) with more tightly controlled audio stimuli: word-final stops are associated with count nouns. The length of the word has no significant effect. In the verbal domain, however, we failed to reproduce the results of Experiment 6 with any audio stimuli. This difference must be due to either the actual presentation of the word (written or audio) or to spurious effects introduced by the precise choice of stimulus items.

In the audio domain, the difference between nouns (where the effect is found) and verbs (where it is not) is currently unexplained by any hypothesis. If the iconic mapping were specific to visually-presented stimuli, then we would expect no effect for either nouns or verbs. If the iconic mapping were generally available across both visual and audio modalities, then we would expect an effect for both nouns and verbs. One potential explanation is that the observed pattern is due to language-specific knowledge. First, because stops and fricatives are already part of subjects' categorical phonological system, existing phonological knowledge may override the effects of iconicity. Second, the effect found for nouns may be driven by English-specific phonotactic knowledge. In particular, in English, the standard plural marker is the fricative */-z/*. Phonotactically, the Obligatory Contour Principle (McCarthy, 1986), may bias language to avoid sequences of fricatives (e.g. "fits" is easier to say than "fifths"). Since only count nouns can be pluralized, this may introduce a general bias against count nouns that end in fricatives. This bias, or another case of non-iconic systematicity in English phonology, may be driving the results observed above for audio stimuli and noun meanings.

Experiment 3

Boundarihood			Divisiveness		
	3a. Verbs	3b. Nouns		3c. Verbs	3d. Nouns
Physical	1.00	1.48	Physical	0.98	0.93
Psych	0.84	0.27	Psych	0.46	0.26

Table 1. General summary of Experiment 3 results. Numerical values provide Cohen's *d* based on z-scores, as an estimate of the effect size. Green squares indicate that count nouns or telic verbs were rated as having a conceptual boundary significantly more than mass nouns or atelic verbs.

SIGN	SPOKEN
Italian Sign Language	Written nonce

		a. Verbs	b. Nouns			a. Verbs	b. Nouns
Exp 1	Physical	1.70	1.94	Exp 6	Physical	0.67	0.71
Exp 2	Psych	1.61	0.51	Exp 7	Psych	0.54	0.20
Nonce gestures				Audio nonce (all physical)			
		a. Verbs	b. Nouns			a. Verbs	b. Nouns
Exp 4	Physical	2.07	0.88	Exp 8	Stop/rep/length	0.25	1.02
Exp 5	Psych	0.99	0.21	Exp 9	Stop/repetition	0.22	1.09
				Exp 10	Stop	0.18	0.67
				Exp 11	Length	0.29	0.23

Table 2. General summary of Experiments and results. Numerical values provide Cohen's d , as an estimate of the effect size. Green squares indicate that stimuli with a boundary (gestural or phonological) received significantly more count/telic responses than stimuli without a boundary. (Experiment 2b showed a significant effect but the the opposite direction.)

General Discussion

The results presented here provide strong and repeated evidence that mapping biases from gesture to meaning are driven by domain general processes. With both natural signs from LIS and highly controlled artificial sign pairs, we have seen that the same mapping bias that exists for verbs also holds for nouns: bounded forms are associated with bounded meanings. The bias is thus not something unique to events; it involves a more general, iconic mapping of abstract properties. Building on this finding, we additionally showed a robust and repeated pattern in which this bias disappears for psychological nouns, which are grammatically categorized as mass or count, but do not have clear spatial or temporal boundaries. We thus conclude that the mapping bias operates on conceptual representations which tend to correlate with grammatical categories as opposed to the grammatical categories themselves.

Explaining the audio results via non-iconic systematicity

Results in the spoken language domain are more nuanced. Written stimuli produced precisely the same interaction as the gestural stimuli, but weaker in effect size. For audio stimuli, we systematically found an effect for noun meanings, but were unable to replicate an effect for verb meanings across several experiments. One of the open questions in the present work is thus why a different pattern appears for audio stimuli. While the effect seen for nouns could be a reflection of the same iconic bias seen for gestures, such an explanation is undermined by the fact that it only appears for nouns, and not for verbs.

At present, the most viable explanation for the audio stimuli is that iconic biases may be overshadowed and neutralized by native language knowledge. Since the difference between /p/ and /f/ is phonologically contrastive in English, categorical perception by English-speaking subjects may disregard the specific acoustic properties relevant to the iconic mapping. Response biases may nevertheless be introduced by non-iconic systematicity specific to English. *Systematicity* describes language-specific statistical regularities between form and meaning that do not have any innate or iconic motivation (Dingemanse et al., 2015). Such regularities (e.g., association of Japanese verbs with coronal consonants /t/, /d/, and /n/) may nevertheless facilitate category classification (Monaghan et al. 2007). In English, we hypothesized that a non-iconic association between count nouns and word-final stop consonants may exist for phonotactic reasons: the plural suffix /z/ is easier to pronounce after stops than after fricatives. This explanation in terms of language-specific knowledge would thus predict that the effect with nouns would disappear for non-English speaking subjects. Conversely, we would expect the iconic bias to reappear for both verbs and nouns if phonological audio stimuli were replaced with non-linguistic stimuli, such as sound effects or musical sequences.

Motivated mappings operating on concepts provide a new window into cognitive categories

The literature reviewed in the paper supports the idea that conceptual categories involved in object and event cognition impose a (non-deterministic) bias on the grammatical categories in language. In the present work, we have seen that these conceptual categories are also involved in an iconic mapping that biases the expected form of verbs and nouns. Interestingly, although the experimental design presented meaning choices as real English words, we found that the iconic bias is not mediated by grammar—it operates directly on conceptual representations. Specifically, we saw that the iconic bias disappears for abstract psychological nouns, even though these may be grammatically encoded as count or mass.

Iconicity and grammar are thus two separate, if interacting, communicative systems, each of which may be subject to cognitive biases. As a result, looking at iconic mappings—especially those acting on abstract logical categories—can provide new evidence for the cognitive underpinnings of language. In the case at hand, our finding confirms that boundariness is a salient conceptual property, and, moreover, that a single abstract property may be shared between events and objects. Adding to existing findings that link telicity and count/mass in language, we have provided a new paradigm, dissociable from grammar, that links boundariness of events and boundariness of objects.

Boundary-marking in sign language

We have suggested that the iconic mapping biases observed here may influence linguistic structure. As discussed above, this appears to be the case for the phonetic representation of telicity on sign language verbs: across many sign languages, gestural boundaries are associated with telic verbs. On the other hand, a number of further questions remain open. Notably, we can

make no claim about the linguistic representation of end-marking in the fully developed grammatical system of native signers. Wilbur (2003; 2008) proposes that end-marking in ASL is in fact a grammatical morpheme that modifies eventive meanings. To fully revisit this claim would require linguistic work with native signers of ASL. In the present work, though, we have provided evidence that the *origins* of this association are not specific to events and are not mediated by grammar.

For sign language nouns, the situation has not been systematically investigated, but the fact that no such tendency of the lexicon has been reported suggests that the effect, if any, may be weaker or more restricted than for sign language verbs. If this is the case, there are several possible explanations. First, we have seen that psychological nouns are not sensitive to the presence of a phonological boundary, which could lead to a weaker effect for nouns in general. Second, the sign stimuli used here manipulated the presence of a temporal boundary, but boundaries can also be represented by using the hands to delineate contours in space. For nouns in sign language, such a space-to-space mapping is likely more accessible than the time-to-space mapping, thus yielding a conflicting phonetic bias for nouns (Meir et al., 2013). In light of the new results presented here, follow-up work can investigate the degree to which we see these mapping biases in other domains—for sign language nouns, as well as for the nouns and verbs of spoken languages.

Higher-level representations: from objects to sets of objects

The present study has restricted itself to lexical properties—morphologically singular nouns and verbs—in order to keep the hypothesis space as small as possible, and reduce additional linguistic confounds. Nevertheless, on the hypothesis of generalized iconicity, one

expects that morphologically complex strategies to mark plurality and aspect should be subject to similar mapping biases as the ones observed here. Notably, a large body of semantic literature shows that mass nouns and plural count nouns share many formal and empirical properties (Link 1983, *i.a.*). Neither the substances in the denotation of *water* nor the collections of objects in the denotation of *pinecones* are single, bounded entities; empirically, both can appear in similar syntactic contexts (*two kilograms of water, two kilograms of pinecones vs. *two kilograms of pinecone*). Indeed, Chierchia (1998) explicitly proposes that mass nouns are in fact lexical items that denote pluralities. The present hypothesis would thus predict that speakers and signers should be biased to associate plural count nouns with non-bounded forms. This, of course, is highly plausible, given the frequency and consistency of iconic derivational forms cross-linguistically. In spoken language, for example, work dating back to Sapir (1921) notes the "self-evident" symbolism of reduplication in order to express nominal and verbal plurality. In sign language, Klima & Bellugi (1979) and many others have documented iconic properties of aspectual modifications, Wilbur, Malaia, & Shay (2012) show that similar modifications can be applied to adjectives, and Schlenker & Lamberton (2019) observe iconic differences in the representation of plural count nouns versus mass nouns in ASL. In all of these cases, morphological complexity adds complexity to the conceptual representations—plural count nouns, for example, while sharing formal properties with mass nouns, also make salient the atomic, bounded parts that comprise the plurality. The present study, alongside current semantic theories, will inform future work on iconic biases in these cases, as well as the path from non-linguistic biases to discrete, grammatical units.

Conclusion

In cognition, spatial and temporal boundaries have been theorized to be central to humans'

perception of objects and events. In language, a related semantic property has characterized the mass/count and telic/atelic oppositions, which have been argued to be grounded in these non-linguistic conceptual representations. Building on observations about sign language, we showed that these conceptual representations are also involved in iconic mapping biases. In a series of experiments, we showed that subjects show an iconic bias to associate bounded forms with bounded meanings, and unbounded forms with unbounded meanings. The representations involved are abstract and domain general: the bias is found both for nouns that denote bounded objects or unbounded substances, and for verbs that denote bounded or unbounded events. We then showed that this mapping bias disappears for psychological nouns (but not psychological verbs), which are grammatically categorized as mass or count, but do not have well defined boundaries in space or time. We conclude that the mapping bias operates on conceptual representations, as opposed to piggybacking on grammatical categories or linguistic paraphrases. The statistical tendency that is observed for sign language thus has its origins in a generally accessible bias involving abstract, iconic mappings that operate on psychological representations.

Supplementary material

<https://osf.io/u5crd/>

Author acknowledgements:

We would like to thank Benjamin Spector, Emmanuel Chemla, Manuel Križ, and Bodo Winter for helpful comments on this work. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grants 324115–FRONTSEM (PI: Schlenker) and 313610–SEMEXP (PI:

Chemla), and ERC H2020 Grant Agreement No. 788077–Orisem (PI: Schlenker). Research was conducted at Institut d’Etudes Cognitives (ENS), which is supported by grants ANR-10-IDEX-0001-02 PSL*, ANR-10-LABX-0087 IEC, and ANR-17-EURE-0017 FrontCog.

References

- Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, 15(1): 5–16.
- Bale, A. C., and Barner, D. (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics*, 26(3):217-252.
- Bennett, M., and Partee, B. (1972). *Toward the logic of tense and aspect in English*. Bloomington: Indiana University Linguistics Club.
- Blasi, D., Wichmann, S., Hammarström, H., Stadler, P., and Christiansen, M. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 113(39):10818–10823.
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., and Spence, C. (2013). “bouba” and “kiki” in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to westerners. *Cognition*, 126:165–172.
- Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. MIT Press, Cambridge, MA.
- Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, 55, 1-5.
- Cheng, C.-Y. (1973). Comments on Moravcsik’s paper. In *Approaches to natural language*, ed. J. Hintikka, J. Moravcsik, and P. Suppes, 286–288. Dordrecht: Reidel.
- Cherries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a principle of object persistence in infancy. *Developmental Science*, 11(3), 427-432.
- Chierchia, G. (1998). Plurality of mass nouns and the notion of “semantic parameter.” In Rothstein, S., editor, *Events and grammar*, pages 53–103. Kluwer, Dordrecht, Netherlands.
- Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese*, 174:99–149.
- Cuskley, C., Simner, J., and Kirby, S. (2015). Phonological and orthographic influences in the bouba-kiki effect. *Psychological Research*, 81(1):119–130.
- Cuxac, C. (2001). French Sign Language: Proposition of a structural explanation by iconicity. In Braffort, A., Gherbi, R., Gibet, S., Teil, D., and Richardson, J., editors, *Gesture-based communication in human-computer interaction*, Lecture Notes in Computer Science, pages 165–184. Springer Berlin Heidelberg.
- Davidson, K., Kocab, A., Sims, A. D., and Wagner, L. (2019). The relationship between verbal form and event structure in sign languages. *Glossa*, to appear.
- Davis, R. (1961). The fitness of names to drawings. a cross-cultural study in Tanganyika. *British*

Journal of Psychology, 52:259–268.

de Saussure, F. (1916). *Cours de linguistique générale*. Payot, Lausanne; Paris.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6(10):654–672.

Dingemanse, M. (2015). Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language*, 39(4):946–970.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., and Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10):603–615.

Dowty, D.R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Reidel.

Gennari, S., Sloman, S., Malt, B., and Fitch, T. (2002). Motion events in language and cognition. *Cognition*, 83:49–79.

Gordon, P. (1985). Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition*, 20(3), 209-242.

Hespos, S., Saylor, M., and Grossman, S. (2009). Infants' ability to parse continuous actions. *Developmental Psychology*, 45(2):575–585.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, 62(2), 169-200.

Inagaki, S., & Barner, D. (2009). Countability in absence of count syntax: Evidence from Japanese quantity judgments. In M. Hirakawa, S. Inagaki, Y. Hirakawa, H. Sirai, S. Arita, H. Morikawa, M. Nakayama, & J. Tsubakita (Eds.), *Studies in Language Sciences (8): Papers from the Eighth Annual Conference of the Japanese Society for Language Sciences*. Tokyo: Kurosio.

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, 41:9–45.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language* 59, 434-446.

Ji, Y. & Papafragou, A. (2020). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition*, 197, 104197.

Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Köhler, W. (1929). *Gestalt psychology*. Liveright, New York, NY.

Krifka, M. (1992). Thematic relations as links between nominal reference and temporal

- constitution. In Sag, I. A. and Szabolcsi, A., editors, *Lexical matters*, pages 29–53. CSLI Publications, Stanford, CA.
- Kuhn, J. and Aristodemo, V. (2017). Pluractionality, iconicity, and scope in French Sign Language. *Semantics and Pragmatics*, 10(6).
- Kulkarni, R., Rothsein, S., and Treves, A. (2013). Statistical investigation into the cross-linguistic distribution of mass and count nouns: Morphosyntactic and semantic perspectives. *Biolinguistics* 7:132–168.
- Lewis, M. L. and Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition*, 153:182–195.
- Liddell, S. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press, Cambridge, UK.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, use and interpretation of language*, pages 303–323. de Gruyter, Berlin, Germany.
- Malaia, E. (2014). It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass*, 8(3):89–98.
- Malaia, E. and Wilbur, R. (2012). Telicity expression in the visual modality. In Demonte, V. and McNally, L., editors, *Telicity, change, and state: A cross-categorical view of event structure*. Oxford University Press, Oxford, UK.
- McCarthy, J. (1986). OCP Effects: Gemination and Antigemination. *Linguistic Inquiry*, 17(2):207–263.
- Meir, I., Padden, C., Aronoff, M., and Sandler, W. (2013). Competing iconicities in the structure of languages. *Cognitive linguistics*, 24(2):309–343.
- Milković, M. (2011). *Verb classes in Croatian Sign Language (HZJ): Syntactic and semantic properties*. PhD thesis, University of Zagreb, Croatia.
- Monaghan, P., Christiansen, M. H., Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive psychology*, 55:259–305.
- Papafragou, A. (2015). The representation of events in language and cognition. In E. Margolis, & S. Laurence (Eds.). *The conceptual mind: New directions in the study of concepts* (pp. 327–346). Cambridge, MA: MIT Press.
- Papafragou, A., Hulbert, A., and Trueswell, J. C. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108:155–184.
- Papafragou, A., Massey, C., and Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of

- motion in thought and language. *Cognition*, 84:189–219.
- Prasada, S., Ferenz, K., and Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition*. 83:141–165.
- Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of Semantics*, 27(3):343–397.
- Samuelson, L., Smith, L. B. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition*, 73:1–33.
- Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12:225–239.
- Schalber K. (2006). Event visibility in Austrian Sign Language (ÖGS). *Sign Language and Linguistics* 9(1/2):207–231.
- Schlenker, P., and Lamberton, J. (2019). Iconic plurality. *Linguistics and Philosophy*, 42(1): 45–108.
- Sharon, T., & Wynn, K. (1998). Individuation of actions from continuous motion. *Psychological Science*, 9(5), 357-362.
- Soja, N. N., Carey, S., and Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, 38:179–211.
- Sonne, T., Kingo, O. S., and Krøjgaard, P. (2016). Occlusions at event boundaries during encoding have a negative effect on infant memory. *Consciousness and Cognition*, 41:72–82.
- Srinivasan, M., & Barner, D. (2020). Lexical, syntactic, and pragmatic sources of countability: An experimental exploration of the mass-count distinction. To appear in *Mass-Count in Linguistics, Philosophy, and Cognitive Science*. Benjamins, Amsterdam.
- Sutton, P. R. and Filip, H. (2016). Mass/count variation: A mereological, two-dimensional semantics. *Baltic International Yearbook of Cognition, Logic and Communication*, 11:1–45.
- Strickland, B. (2016). Language reflects "core" cognition: A new hypothesis about the origins of cross linguistic regularities. *Cognitive Science*, 41: 70–101
- Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., and Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences*, 112(19):5968–5973.

- Swallow, K., Zacks, J., & Abrams, R. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology*, 138, 236–257.
- Tenny, C. (1994). *Aspectual roles and the syntax-semantic interface*. Dordrecht: Kluwer Academic.
- Thompson, P. and Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology*, 64:2392–2404.
- Ulan, R. (1978). Size-sound symbolism. In Greenberg, J. H., Ferguson, C. A., and Moravcsik, E. A., editors, *Universals of human language, Vol. 2: Phonology*, pages 527–568. Stanford University Press, Stanford, CA.
- vanMarle, K. and Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, 14(5):498–504.
- vanMarle, K. and Wynn, K. (2011). Tracking and quantifying objects and non-cohesive substances. *Developmental Science*, 14(3):502–515.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review* 66(2): 143–160.
- Westermann, D. (1927). Laut, ton und sinn in westafrikanischen sudansprachen. In Boas, F., editor, *Festschrift Meinhof*, pages 315–328. L. Friederichsen, Hamburg, Germany.
- Wilbur, R. (2003). Representations of telicity in ASL. In *Chicago Linguistic Society 39*, pages 354–368.
- Wilbur, R. (2008). Complex predicates involving events, time and aspect: Is this why sign languages look so similar? In Quer, J., editor, *Theoretical Issues in Sign Language Research*, pages 217–250, Hamburg, Germany. Signum Press.
- Wilbur, R., Malaia, E., and Shay, R.E. (2012). Degree Modification and Intensification in American Sign Language Adjectives. In Aloni, Kimmelman, Roelofsen, Sassoon, Schulz, and Westera (eds.), *Logic, Language, and Meaning: 18th Amsterdam Colloquium*. Amsterdam: Springer.
- Zacks, J. M., Kumar, S., Abrams, R. A., and Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, 112:201–216.
- Zacks, J., Speer, N., Swallow, K., Braver, T., and Reynolds, J. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2):273–293.
- Zacks, J. and Swallow, K. (2007). Event segmentation. *Current Directions in Psychological Science*, 16(2), 80–84.

Appendix A

The following stimuli and meaning choices were used.

LIS stop stimuli (Experiments 1 and 2):

youtu.be/BX4RgBrFL2c (ENTER), youtu.be/uUt_BFMUswM (SELL),
 youtu.be/lWy8flaGQWM (FORGET), youtu.be/aNE_bM12ggQ (LEAVE),
 youtu.be/fOIpVoq26rE (DIE), youtu.be/0xKOAmFN2Rw (BUY),
 youtu.be/JKQfPdENKNE (DECIDE), youtu.be/Pq1INx1_L9M (CONFIRM),
 youtu.be/FNRNTyrj-lc (MARRY)

LIS non-stop stimuli (Experiments 1 and 2):

youtu.be/VuePBBSySFk (RUN), youtu.be/0duaBmskr4w (PLAY),
 youtu.be/eYYEmTJ0xd8 (NEGOTIATE), youtu.be/hPQDdHW9Deg (FLOAT),
 youtu.be/Gs0RNUgwezEE (TALK), youtu.be/L3SejJKVs5E (DISCUSS),
 youtu.be/jWMM_sa7PNM (IMAGINE), youtu.be/zqBaeY45ZU (THINK),
 youtu.be/O3Q2cbEckYk (PONDER)

Nonce sign stop stimuli (Experiments 4 and 5):

youtu.be/WfGO5m2Xmz0, youtu.be/_DZWJob5BMw, youtu.be/OpwGub42Puo,
 youtu.be/taCCyRol_S0, youtu.be/YXsoYA1-R0Q, youtu.be/zv2j6QMJTiq,
 youtu.be/wD3fSdA36xc, youtu.be/jrRA1FYhjFU, youtu.be/LHlegClffnk

Nonce sign non-stop stimuli (Experiments 4 and 5):

youtu.be/WxiQvFpW3DA, youtu.be/aTsa6glY7HE, youtu.be/nxTmcSUQKHs,

youtu.be/l6bqwHB4jQw, youtu.be/8-JN08KGeMY, youtu.be/EGLCvfjCgXQ,
youtu.be/yx9WU08LDe0, youtu.be/PPedOlGl5r8, youtu.be/7y76phC0hxQ

Nonce word stop written stimuli (Experiments 6 and 7):

zod, rab, sab, bip, cip, fot, quot, tot, lod

Nonce word non-stop stimuli (Experiments 6 and 7):

fossoss, razzazz, bifif, tossoss, lovov, sazzazz, cifif, zovov, quossoss

Nonce word short stop audio stimuli (Experiments 8, 10, and 11):

/gop/, /lop/, /ret/, /pot/, /tok/, /mek/, /leb/, /rod/, /mag/

Nonce word long stop audio stimuli (Experiments 9 and 11):

/kalop/, /lærip/, /rɛmæt/, /porat/, /timok/, /molek/, /lorib/, /ramud/, /milag/

Nonce word short non-stop audio stimuli (Experiment 10):

/gof/, /lof/, /res/, /pos/, /toʃ/, /mɪʃ/, /lev/, /roz/, /maz/

Nonce word long non-stop audio stimuli (Experiments 8 and 9):

/pofof/, /mæfæf/, /lɪsɪs/, /tasas/, /kʊʃʊʃ/, /rɛʃɛʃ/, /mɪvɪv/, /lozoz/, /razaz/

Physical verb meaning choices (Experiments 1, 4, 6, 8, 9, 10, and 11):

stroll/fall, drip/start, skate/quit, glide/catch, flow/throw, walk/break, sparkle/shatter,
tremble/settle, limp/slip

Physical noun meaning choices (Experiments 1, 4, 6, 8, 9, 10, and 11):

rain/coin, powder/apple, fog/bag, moisture/pinecone, grass/knife, air/ball,
weather/window, water/needle, sand/rock

Psychological verb meaning choices (Experiments 2, 5, and 7):

await/conclude, dream/choose, consider/discover, believe/concede,
contemplate/designate, observe/notice, desire/resolve, visualize/identify, dislike/select

Psychological noun meaning choices (Experiments 2 and 7):

knowledge/notion, contempt/concept, attention/decision, bravery/idea, sympathy/theory,
comprehension/realization, faith/choice, certainty/objective, intelligence/preference

Psychological noun meaning choices (Experiment 5):

fun/mood, information/decision, attention/opinion, knowledge/concept, advice/theory,
anger/choice, wisdom/notion, bravery/idea, intelligence/objective

Appendix B

Following are the model summaries for the maximal models with observed interactions.

Experiment 3a and 3b (boundariness ratings of conceptual representations)

Formula: $\text{Score} \sim \text{Category} * \text{Domain} * \text{PoS} + (1 | \text{Participant}) + (1 | \text{Word})$

AIC = 4279.36; Pseudo- R^2 (Nagelkerke, Cragg and Uhler) = 0.404708

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
Categorymass	-2.7106	0.2642	-10.258	< 2e-16	***
Domainpsych	-2.8441	0.2569	-11.071	< 2e-16	***
PoSverb	-1.1055	0.5529	-1.999	0.0456	*
Categorymass:Domainpsych	2.8966	0.3225	8.982	< 2e-16	***
Categorymass:PoSverb	1.6731	0.3286	5.091	3.55e-07	***
Domainpsychological:PoSverb	2.4875	0.3264	7.622	2.50e-14	***
Categorymass:Domainpsych:PoSverb	-2.4587	0.4278	-5.748	9.03e-09	***

Experiment 3c and 3d (divisiveness ratings of conceptual representations)

Formula: $\text{Score} \sim \text{Category} * \text{Domain} * \text{PoS} + (1 | \text{Participant}) + (1 | \text{Word})$

AIC = 4213.74; Pseudo- R^2 (Nagelkerke, Cragg and Uhler) = 0.2879700

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
Categorymass	1.3116	0.2266	5.787	7.16e-09	***
Domainpsych	1.1783	0.2074	5.681	1.34e-08	***
PoSverb	0.6153	0.4832	1.273	0.202870	
Categorymass:Domainpsych	-1.1162	0.2869	-3.891	9.97e-05	***
Categorymass:PoSverb	-0.1645	0.3045	-0.540	0.589036	
Domainpsychological:PoSverb	-1.0051	0.2881	-3.489	0.000485	***
Categorymass:Domainpsych:PoSverb	0.6834	0.4061	1.683	0.092416	.

Experiments 1 and 2 (LIS signs)

Formula: `Telic.count.response ~ Boundary * Domain * PoS + (1 | Participant) + (1 | Meanings) + (1 | Stimulus)`

AIC = 9177.4; Pseudo- R^2 (Nagelkerke, Cragg and Uhler) = 0.1251880

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.3526	0.1426	2.473	0.0134	*
Boundaryno-boundary	-1.3893	0.1668	-8.329	< 2e-16	***
Domainpsych	-0.2589	0.1528	-1.694	0.0902	.
PoSverb	0.2327	0.1550	1.501	0.1332	
Boundaryno-boundary:Domainpsych	1.7809	0.1422	12.526	< 2e-16	***
Boundaryno-boundary:PoSverb	0.1775	0.1458	1.217	0.2235	

Domainpsych:PoSverb	0.2914	0.2185	1.334	0.1823
Boundaryno-boundary:Domainpsych:PoSverb	-1.6441	0.2023	-8.125	4.46e-16 ***

Experiments 4 and 5 (Nonce signs)

Formula: Telic.count.response ~ Boundary * Domain * PoS + (1 | Participant) + (1 | Meanings) + (1 | Stimulus)

AIC = 7574.8; Pseudo- R^2 (Nagelkerke, Cragg and Uhler) = 0.0901661

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.09563	0.15642	0.611	0.540960
Boundaryno-boundary	-0.61065	0.15802	-3.864	0.000111 ***
Domainpsych	0.16411	0.18995	0.864	0.387610
PoSverb	0.81543	0.18979	4.296	1.74e-05 ***
Boundaryno-boundary:Domainpsych	0.47602	0.15434	3.084	0.002041 **
Boundaryno-boundary:PoSverb	-0.74494	0.15143	-4.919	8.68e-07 ***
Domainpsych:PoSverb	-0.62152	0.27121	-2.292	0.021927 *
Boundaryno-boundary:Domainpsych:PoSverb	0.18255	0.22160	0.824	0.410071

Experiment 6 and 7 (Written nonce words)

Formula: Telic.count.response ~ Boundary * Domain * PoS + (1 | Participant) + (1 | Meanings) + (1 | Stimulus)

AIC = 10430.3; Pseudo- R^2 (Nagelkerke, Cragg and Uhler) = 0.0199022

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.31525	0.09196	3.428	0.000608	***
Boundaryno-boundary	-0.53008	0.12093	-4.383	1.17e-05	***
Domainpsych	-0.05754	0.10393	-0.554	0.579838	
PoSverb	-0.05790	0.10077	-0.575	0.565564	
Boundaryno-boundary:Domainpsych	0.39955	0.13588	2.941	0.003276	**
Boundaryno-boundary:PoSverb	0.02096	0.13760	0.152	0.878900	
Domainpsych:PoSverb	0.02126	0.14280	0.149	0.881641	
Boundaryno-boundary:Domainpsych:PoSverb	-0.29550	0.18603	-1.588	0.112179	