Encoding individuals in language using syntax, words, and pragmatic inference

https://escholarship.org/uc/item/569855p4

Authors
Srinivasan, Mahesh
Barner, David

Publication Date
2016

Data Availability
The data associated with this publication are available upon request.

Peer reviewed
Advanced Review

Encoding individuals in language using syntax, words, and pragmatic inference

Mahesh Srinivasan1* and David Barner2

How does linguistic structure relate to how we construe reality? In many languages, countable individuals like objects are typically labeled by count nouns (e.g., two rabbits, every truck, etc.), while unindividuated masses like substances are typically labeled by mass nouns (e.g., much mud, barrel of oil, etc.) (Quine WVO. Word and Object. Cambridge, MA: MIT Press; 1960). These facts have led researchers to propose that learning mass–count syntax affects how speakers perceive objects and substances or alternatively that an understanding of this distinction—or one between individuals and nonindividuals—scaffolds the acquisition of mass and count nouns. Here, we evaluate these ideas and describe how recent developments in the literature have fundamentally changed our understanding of the mass–count distinction and how it relates to individuation. Across three sections, we show that a simple distinction between countable individuals and nonindividuals cannot provide a foundation for the mass–count distinction (e.g., because many mass nouns like furniture and luggage can denote individuals). Furthermore, we show that mass–count syntax does not shape whether items are construed as individuals or not, but instead allows speakers to select from a set of universally available meanings (e.g., because speakers of all languages quantify objects and substances similarly). We argue that a complete understanding of how mass–count syntax encodes reality requires understanding how different aspects of language—syntax, lexical roots, word meanings, and pragmatic inference—interact to encode abstract, countable individuals. © 2016 Wiley Periodicals, Inc.

INTRODUCTION

Some entities, such as rabbits, shoes, and trucks, are typically construed by humans as individuals. For example, we have a clear idea about where one rabbit ends and another begins, allowing us to count rabbits. By contrast, other entities, like water, dust, and mud, are not typically construed as individuals. We can only count substances like water or mud if we first specify units for counting, like bottles or piles. This distinction between individuated and nonindividuated phenomena does not reduce to perception, but is instead conceptual: We are capable of construing the very same physical entity—e.g., a glass window—either as a kind of individual (a window) or not (some glass). Clearly, then, language plays some role in affecting how speakers construe the world.

In this study, we explore how conceptual content relates to linguistic form by considering how individuated and nonindividuated phenomena are represented by different linguistic structures. For
example, in English, countable things are often labeled using nouns that appear in count syntax: Nouns that can be singular or plural (e.g., *rabbit* vs *rabbits*), and modified by numerals (e.g., *three rabbits*) and quasi-cardinal quantifiers like *many, these, several, and every*. In contrast, uncountable phenomena are typically labeled using nouns that appear in mass syntax: These nouns cannot be pluralized (*muds*) and cannot be directly modified by numerals (e.g., *three muds*) or quasi-cardinal quantifiers (*many muds*), though they can appear with quantifiers like *little or much*. These observations have led some to propose a direct correspondence between count nouns and individuated phenomena like objects, on the one hand, and mass nouns and nonindividuated phenomena like substances, on the other hand.1 Furthermore, by some accounts, our ability to construe the referents of nouns as individual objects or nonindividuated substances may itself be shaped by learning mass–count syntax.2–4 While by other accounts, an ontological object–substance distinction helps children learn mass and count nouns in the first place.11,12 Here, we review previous work on the mass–count distinction and argue that, while many mass and count nouns denote objects and substances in the world, a focus on this distinction is ultimately misleading. Instead, we argue that the mass–count distinction must instead be captured at a more abstract level that accounts for how different aspects of language—syntax, lexical roots, and concepts—encode abstract, countable individuals.

Benjamin Whorf5 famously hypothesized that language shapes how we conceptualize the world, arguing that ‘...language is not simply a reporting device for experience but a defining framework for it.’ Applying this idea to the case of individuation, Whorf suggested that the distinction between objects and substances made by speakers of Indo-European languages is in fact imposed by their language. By this reasoning, an ontological object–substance distinction may not be shared by speakers of languages that lack a mass–count distinction, and children who have not yet acquired mass–count syntax may not be able to conceptualize entities as objects, but only as undifferentiated portions of experience: e.g., as ‘undetached rabbit parts’ or ‘time slices of rabbit’ as opposed to ‘a rabbit.’11

Although this claim—that mass–count syntax shapes cognition—is controversial, many scholars have accepted a premise of this argument—i.e., that there exists a systematic correspondence between mass–count syntax and individuation. For example, in his influential writings, Quine1 argued that while count nouns ‘possess built in modes, however arbitrary, of dividing their reference,’ mass nouns label various kinds of nonindividuals. By his account, while quantities corresponding to count nouns like *chair* are determined by enumerating units provided by the noun (e.g., chair-sized things), mass nouns like *clay* do not provide units of quantification and must be quantified according to continuous dimensions like mass or volume. Although many scholars have rejected the strong Whorfian hypothesis, they have nevertheless accepted Quine’s semantic proposal,5–14 and have argued that a semantic distinction between objects and substances—or a more abstract distinction between individuals and nonindividuals—helps children bootstrap into mass–count syntax.

The past decades have seen a flourishing of research on grammatical number and the mass–count distinction, which has led to a rapid accumulation of new facts across different languages and developmental periods.15–26 This review describes how these recent developments have fundamentally changed our understanding of how the mass–count distinction relates to individuation, from a focus on the object–substance distinction to an approach in which mass–count syntax is viewed as a system for expressing modes of measurement and quantification. In what follows, we first review studies that evaluate the Quinean hypothesis that while count nouns correspond to individuals, mass nouns cannot. Against this hypothesis, we show that count syntax is not the only factor that affects individuation, and thus that there is not a perfect correspondence between mass–count syntax and the object–substance distinction (e.g., we show that even mass nouns like *furniture* and *luggage* can denote countable things). Second, we evaluate the Whorfian prediction that speakers of classifier languages, which lack count syntax, should individuate entities differently than speakers of mass–count languages. We review evidence showing that speakers of all languages quantify similarly, and argue that lexical concepts provide a universal source of individuation that is independent of mass–count syntax. Thus, syntax allows speakers to select from—rather than create—alternate construals of the world. Finally, we consider how lexical concepts specify which countable individuals a word refers to—e.g., how *chair* picks out only chairs. We argue that such information is not fully supplied by concepts but instead depends on our pragmatic ability to contrast word meanings with one another. In sum, we will suggest that a full understanding of how language represents countable individuals requires considering how syntactic, lexical, conceptual, and pragmatic processes interact.
THE SEMANTIC BASIS OF THE MASS–COUNT DISTINCTION: SYNTACTIC AND LEXICAL SOURCES OF INDIVIDUATION

Quine’s proposal, that count nouns denote countable individuals and mass nouns denote nonindividuals, carries intuitive appeal. For one, many of the mass and count nouns that first come to mind follow this generalization: While count nouns like cat and chair denote canonical objects, mass nouns like mud and sand denote canonical substances. Syntax also has a demonstrable effect on interpretation, because when countable nouns are used in mass syntax, the coerced meanings tend to label nonindividuated substances (e.g., there is banana all over the floor). Furthermore, linguistic tests point to differences in the part-whole relations encoded by mass and count nouns, at least on first glance. Quine, e.g., claimed that only mass nouns refer cumulatively: If X is sand and Y is sand, then X and Y taken together is also sand, but the same is not true for a count noun like chair. Cheng27 later proposed a second distinction between count and mass nouns, and claimed that because only mass nouns lack minimal units for quantification, they are subject to divisity: When some sand is divided into two smaller portions, the two portions are each sand, but when a chair divided into two arbitrary portions, the two portions are not chairs (for review, see Ref 28).

This proposed correspondence between mass–count syntax and individuation has been widely adopted, and has led to different accounts of how acquiring mass–count syntax relates to our cognitive ability to distinguish objects from substances. Quine famously argued that the object–substance distinction is a cultural construction, and that young infants do not represent the world in terms of stable objects, but instead as undifferentiated portions of experience; This argument, coupled with the idea that count nouns denote individuals like objects, has contributed to the Whorfian thesis that acquiring mass–count syntax leads children to understand the object–substance distinction. However, many developmental psychologists have argued against this idea, and have instead proposed that young children’s acquisition of count and mass nouns is itself scaffolded by a prior understanding of the object–substance distinction. By this account, the first count and mass nouns children learn correspond to objects and substances, respectively, with other nouns assigned count or mass status when they appear in similar distributional profiles to already-acquired nouns. To test these ideas, studies have explored both whether children’s understanding of the object–substance distinction emerges prior to acquiring mass–count syntax, and whether children form mappings between objects and count nouns on the one hand, and substances and mass nouns on the other hand.

Contrary to the Whorfian hypothesis, a large body of evidence suggests that children represent an object–substance distinction well before acquiring mass–count syntax. To begin, even prelinguistic infants appear to have a cursory object concept,29,30 which allows them to individuate objects and trace their numerical identity. Infants also quantify objects differently from nonsolid substances, suggesting an early knowledge of objects and substances.31 Furthermore, toddlers appear to rely on this knowledge when learning new words: Prior to acquiring mass–count syntax, 2-year olds learning English extend new nouns differently if the noun first labels an object (preferring to extend the label to other items of the same shape) than if it first labels a nonsolid substance (preferring to extend the label to other items of the same substance24). Similar findings have also been reported for children who have not acquired mass–count syntax: Like English-learning children, Japanese-learning 2-year olds distinguish solid objects from nonsolid substances when extending new nouns, as do children learning Mandarin Chinese. Together, these studies suggest that acquiring mass–count syntax does not lead children to make an object–substance distinction.

Could an understanding of the object–substance distinction instead provide a foundation for learning mass and count nouns? If it does, then children might initially assume that labels for objects are count nouns, and that labels for substances are mass nouns. Contrary to this prediction, Gordon provided evidence that young children do not categorize nouns as mass or count simply on the basis of whether they label objects or substances, but instead prefer to classify nouns on the basis of syntactic cues. In his study, 3- to 5-year olds heard either a singular count noun (This is a garn) or mass noun (This is some garn) used to label one of three kinds of training stimuli: a portion of nonsolid material (in a test tube), a solid object, or a collection of solid objects. Gordon then showed children a collection of the objects (if children had seen a solid object or a collection in training) or a set of test tubes of the nonsolid stuff (if children had seen the single test tube of stuff in training), and asked whether children would pluralize the word—which would suggest that they categorized it as a count noun—by having them complete the phrase, ‘Over here we have more …. ’ Interestingly, children’s pluralization was more affected by the initial syntactic context in which the
novel words had been introduced than by the ontological status of their referents. Thus, children tended to pluralize the novel words when they were first introduced in count syntax, even if they had initially labeled substances; children also avoided pluralizing the words when they were first introduced in mass syntax, even if they had initially labeled objects.

The findings of Gordon and others suggest that a simple ontological distinction between objects and substances does not provide a foundation for the mass–count distinction (an idea also articulated by many linguists). Instead, Gordon proposed that mass and count categories have a more abstract foundation, and ‘...are defined in terms of grammatical roles which include their proper quantificational functions’ (Ref 6, p. 211). Also arguing for this analysis, Bloom—echoing an earlier proposal by Quine—proposed that children infer that a word is a count noun if its referent has been construed as an individual (e.g., as an object, unified collection of objects, portion of substance, sound, etc.), and infer that it is a mass noun if its referent has been construed as unindividuated (see also Ref 14). On Bloom’s account, such syntax–semantics mappings are bidirectional, such that the application of count syntax leads children to construe even substances as individuals (when they can be quantified as individual portions), and mass syntax leads collections of objects to be construed as nonindividuals, explaining Gordon’s findings. Providing credence to the idea that an individual/nonindividual distinction could provide a foundation for bootstrapping count and mass nouns, Bloom noted that infants’ conceptions of individuals are not limited to physical objects, but are more abstract and encoded in early linguistic and counting abilities.

However, although infants may have an abstract understanding of individuation from early in acquisition, it is unclear whether there is a one-to-one correspondence between count and mass nouns and individuals and nonindividuals, as argued by Quine, Bloom, and others. While linguistic tests of cumulativity and divisibility have been used to argue that only count nouns individuate because mass nouns refer cumulatively and are subject to divisity (and thus do not have minimal parts), these generalizations are subject to important exceptions, as argued by Gillon. For example, although mass nouns like sand refer cumulatively—if A is sand and B is sand, then A and B together is sand—so too do plural count nouns like chairs. Furthermore, although many mass nouns are subject to divisity—a portion of sand divided in multiple portions is still sand—this is not true of other mass nouns like furniture, luggage, and jewelry, while it is true of some count nouns like string, fence, and rope (for additional discussion, see Ref 23, 48).

To account for the fact that such linguistic tests do not perfectly differentiate count and mass nouns, some linguists have argued against the Quinian account that individuation is restricted to count nouns and have instead proposed that mass nouns can also have minimal parts and thus individuate. For example, Gillon proposed that while count nouns always denote countable individuals, mass nouns are unspecified, and can either denote individuals or nonindividuals (see also Chierchia, for a similar but stronger proposal that all mass nouns denote pluralities of individuals). By these accounts, the best indication of a particular noun’s denotation—e.g., whether it individuates or not—is the physical properties of the things in the world that the noun refers to. Thus, furniture individuates because furniture in the world has minimal parts, like chairs and tables. Furthermore, flexible words that can appear both as count and mass nouns—e.g., string, stone, rope—are argued to have the same meanings and to denote individuals in both their count and mass forms, since their ability to appear in count syntax requires that their referents in the world have minimal parts, and because these words have the same referents when appearing in mass syntax.

According to Quinian accounts, however, the fact that linguistic tests do not distinguish mass nouns like furniture and clothing from count nouns like chair and ball need not imply that such mass nouns denote countable individuals. In particular, although these words label referents that have minimal parts in the world, these referents may not be construed as individuals when labeled by mass nouns; These accounts thus argue that the denotation of a word is not determined by the physical attributes of its referent but instead by how it is mentally represented. Articulating this idea, Bloom argued that ‘mappings relevant to the study of linguistic competence must be between grammatical classes and cognitive classes—not classes of entities in the world’ (Ref 5, p. 45). For example, mass nouns like furniture or waterfowl may not denote individuals, because they could cause speakers to construe entities as unindividuated masses—e.g., and conceptualize ‘a swan, several ducks, and a heron on a lake as an unindividuated group called waterfowl...’—thus preserving the Quinian hypothesis that only count nouns denote individuals. Furthermore, flexible words like string, stone, and rope could have different denotations when used in count syntax (where
they lead to individual construals), and in mass syntax (where they lead to nonindividual construals). This position is clearly at odds with that of Gil- lon\textsuperscript{28,33} and Chierchia\textsuperscript{34,44}, who argue that a word’s denotation with its referent, and thus argue that flexible nouns must have the same denotations when appearing in count or mass syntax. For example, Chierchia\textsuperscript{34} argued that the English mass noun *hair* and its translation-equivalent in Italian, *capelli*—a count noun—must have the same denotations: ‘…on most theories, Pavarotti’s hair is some kind of atom-less substance in English, but turns into an atomic one in Italian. If we do not want semantics to start looking like magic, we have to say that in the real world ‘hair’ and ‘capello’ obviously denote the same stuff’ (Ref 34, p. 88).

As can be seen, the two accounts of the mass–count distinction described above disagree both about whether mass nouns like *furniture* denote individuals, and also about whether flexible nouns like *string* have different denotations in count and mass syntax. What is needed, then, is an experimental measure of construal that confirms the predictions that all previous accounts make about uncontroversial cases—e.g., that object-count nouns like *chair* and *shoe* will denote individuals and that substance-mass nouns like *water* and *oil* will denote nonindividuals—and that can also decide between the different predictions these accounts make about object-mass nouns like *furniture* and *clothing*, and flexible nouns like *string* and *stone*.

To address this challenge, Barner and Snedeker\textsuperscript{17} used a quantity judgment task, in which 4-year olds and adults were introduced to two characters, where one character had a large object or part of substance (e.g., a giant shoe, or a portion of mustard), and the other had three tiny objects or portions (e.g., three tiny shoes or portions of mustard), and were then asked which character had more (e.g., Who has more shoes/mustard?). Using this method, Barner and Snedeker\textsuperscript{17} tested subjects’ interpretation of uncontroversial cases like substance–mass nouns (e.g., *mustard* and *ketchup*) and object-denoting count nouns (e.g., *shoe* and *candle*), as well as their interpretation of more controversial cases like object–mass nouns (e.g., *furniture* and *jewelry*), and flexible nouns (e.g., *string* and *stone*), where the latter were presented to some subjects in mass syntax, and to others in count syntax. The findings were as expected for uncontroversial cases: subjects based their judgments on number for object-denoting count nouns (e.g., indicating that three tiny shoes are *more shoes* than one big one), but on mass or volume for substance–mass nouns (e.g., indicating that a large portion of *mustard* is *more mustard* than three tiny portions). Critically, however, subjects of all ages also based their judgments on number for object–mass nouns like *furniture* and *jewelry*. These results are consistent with the predictions of Gil- lon\textsuperscript{28,33} and Chierchia,\textsuperscript{34,44} and are difficult to explain for Quin-nian accounts, which claim that no mass nouns denote individuals.\textsuperscript{1,6,7,10}

Results from flexible nouns like *string* and *stone* provided important additional data. Recall that on the Quinian view, an empirical measure is only relevant if it measures how speakers construe referents, and not just how those referents appear in the world. Consistent with the idea that the quantity judgment task is sensitive to speakers’ construal of referents, Barner and Snedeker\textsuperscript{17} found that subjects shifted their judgments of flexible nouns like *string* and *stone* based on whether such nouns were presented in count or mass syntax. Specifically, when presented in count syntax (e.g., Who has *more stones?*) subjects made their judgments according to number (suggesting that they construed the referents as individuals), but instead based their judgments on mass or volume when the nouns were presented in mass syntax (e.g., Who has *more stone?). Critically, these results show that the quantity judgment task is sensitive to construal, and thus that when subjects judged object–mass nouns like *furniture* according to number, this reflected how speakers construed the referents of mass nouns, providing evidence against Quinian accounts. However, the results from flexible nouns are also difficult to explain for the accounts of Gil- lon and Chierchia, because they show that whether or not a noun individuates is not determined simply by whether minimal parts of referents can be found in the world; instead, flexible nouns have different meanings when they appear in count or mass syntax.\textsuperscript{2}

Taken together, the results from Barner and Snedeker\textsuperscript{17} suggest an asymmetric relationship between the mass–count distinction and individuation: While count nouns always denote individuals and quantify by number, mass nouns can denote either individuals or nonindividuals, and quantify by number, volume, mass, or other continuous dimensions. Furthermore, in contrast to the accounts of Gil- lon and Chierchia, the denotations of nouns are not determined by physical attributes of referents in the world, but instead correspond to how those referents are construed, such that for cases like *stone* and *string*, a particular entity can be construed either as an individual or not. Interestingly, this generalization—that count nouns denote individuals, but mass nouns can denote a variety of
phenomena—applies not only to object and substance nouns, but also to action-denoting nouns: When verbs labeling continuous actions like _dance or walk_ are used as mass nouns, subjects base quantity judgments on dimensions like time or distance traveled (e.g., Who did more _dancing_?), whereas when they are used in count syntax judgments are based on the total number of discrete actions (e.g., Who did more _dances_?). However, noncontinuous, iterative actions like _jump and kick_ always quantify by number, regardless of whether they are used in mass syntax or count syntax. Thus, much like _furniture and luggage_, subjects base judgments of ‘Who did more _jumping_?’ on number, despite the fact that in this context _jumping_ is a mass noun (for further discussion of event quantification and the mass–count distinction, see Refs 52–55).

Findings from word extension tasks also support the idea of an asymmetry between count and mass nouns. In particular, studies have repeatedly found that although children assume that count nouns label individuals, they allow mass nouns to label either individuals or nonindividuals. For example, Soja taught children a novel count or mass noun either for a simple solid object (e.g., a pyramid made out of wood and molding clay) or a nonsolid substance (e.g., a backward-S shaped portion of Nivea cream), and then asked subjects to extend the label (e.g., ‘find the blicket’) to either an entity that matched in shape but differed in material (indicating an object construal) or one that matched in material but differed in shape. Strikingly, 2.5-year olds often extended count nouns by shape, not just when they had labeled solid objects (90% of the time) but even when they had labeled substances (49% of the time). In contrast, 2.5-year olds did not show a strong substance bias for mass nouns, and extended mass nouns for solid objects by substance only 24% of the time (see also Ref 24). Subrahmanyan and colleagues found a similar pattern, showing that 3-year olds extended both count nouns and mass nouns for novel objects by shape (90 and 86%, respectively). Although this shape bias for mass nouns declined with age, even adults were willing to extend mass nouns by shape on 30% of trials. The tendency to extend mass nouns by shape also increases when the objects are complex, suggesting that they perform a function: for such objects, Barner and Snedeker found that both 3-year olds and adults extended novel mass nouns by shape more than half of the time. In sum, subjects almost always assume that count nouns for novel objects label individuals and thus extend them by shape, but do not make this assumption for mass nouns, and extend them either by substance or shape.

To explain this asymmetric relationship between mass–count syntax and individuation, researchers have argued that syntax is not the only source of individuation, and instead that syntactic and lexical processes interact to encode individuals. By this account, count syntax takes lexical representations that are not prespecified for individuation as inputs, and returns individuated representations, ensuring that count nouns will quantify by number, whereas mass syntax applies an identity function to whatever lexical representation it takes as an input, allowing mass nouns to quantify along various dimensions. Thus, mass nouns like _furniture and clothing_ denote individuals and quantify by number because they are prespecified for individuation in the lexicon, and cannot appear as count nouns because count syntax requires unindividuated lexical representations (e.g., *two furnitures*). Also by this reasoning, flexible nouns like _stone and string_ and coerced nouns like ‘a _beer_’ are not prespecified for individuation, such that they can appear both in count syntax (where they denote individuals due to count syntax) and in mass syntax (where they do not receive an individuating function from syntax, and thus take on an unindividuated interpretation).

Importantly, according to this model, a noun’s lexical representation—which can be prespecified for individuation, in the case of mass nouns—is distinct from the concept that the noun is linked to. For example, while the lexical representation of a count noun like _chair_ may not be prespecified for individuation—allowing it to receive the individuating function of count syntax—it is still linked to the concept CHAIR, which provides information about which individuals in the world count as instances of chairs. Thus, although count syntax may generate the intuition that a particular noun like _chair_ will label individuals, concepts will constrain the precise interpretation that the noun receives: e.g., such that _chair_ is linked to specific individuals like whole chairs as opposed to other objects. Similarly, to interpret flexible words like _string and stone_ as count nouns—or coerced nouns like ‘a _beer_’—we have to identify the relevant individuals in the world, and this process may be mediated by concepts (and the availability of individuals in the context). In the next section, we consider whether concepts can provide a source of individuation for nouns in languages that lack mass–count syntax, leading speakers of all languages to quantify similarly, or if instead learning mass–count syntax...
transforms how speakers construe entities as individuals or nonindividuals.

**DOES LEARNING MASS–COUNT SYNTAX TRANSFORM NOUN MEANINGS? EVALUATING THE WHORFIAN HYPOTHESIS**

In the previous section, we reviewed evidence against the Whorfian proposal that learning count syntax enables children to conceptualize entities as individuals as opposed to nonindividuated masses. However, some research has suggested that in classifier languages like Japanese and Yucatec Mayan—which lack mass-count syntax—nouns have fundamentally different meanings than they do in languages with mass-count syntax like English. By these accounts, while count nouns in languages like English provide built-in minimal units for quantification, nouns in classifier languages have nonindividuated denotations by default, and can only denote individuals when combined with classifiers. For example, according to Lucy, the Yucatec noun for banana (ba’as) can have various different meanings including ‘banana-fruit,’ ‘banana-leaf,’ ‘banana-tree,’ and ‘banana-bunch’ depending on which Yucatec classifier it is used with. Building on the idea that count nouns in languages with mass-count syntax provide built-in units while Yucatec nouns do not, Lucy proposed that English count nouns draw attention to the shape of a referent—consistent with an individuated interpretation—while Yucatec nouns draw attention to the referent’s material composition—consistent with a nonindividuated interpretation.

Supporting the proposal of Lucy, several studies suggest that although children learning classifier languages initially interpret novel words for objects and substances in similar ways to children learning English, they diverge later in life, such that English-learning children are more likely to extend new words by shape than children learning Japanese or Yucatec. For example, using a word extension task, Imai and Gentner taught English and Japanese-speaking adults and children a novel word for either a nonsolid (e.g., an omega-shaped portion of Nivea cream), a solid that had a simple shape (e.g., a cork half pyramid), or a solid that had a complex shape (e.g., an apple corer). Subjects were then asked to extend the label to either an entity that matched in shape but not material, or one that matched in material but not shape. Interestingly, from around 2.5 years of age, Japanese speakers were more likely to extend the noun according to material than English speakers.

To explain their findings, Imai and Gentner argued that entities can be conceptualized as falling along an individuation continuum and that speakers of mass-count and classifier languages may draw different boundaries between individuals and nonindividuals. For example, by virtue of learning count syntax, English speakers may be particularly sensitive to individuable things, such that they construe a wider set of entities as individuals than do speakers of classifier languages like Japanese. Thus, by this view, speakers of different languages do not draw on a universal set of meanings; instead, speakers of languages with mass-count syntax are more likely to have nouns that individuate than speakers of classifier languages.

However, others have argued that cross-linguistic differences found in novel word extension tasks may not reveal genuine effects of language on thought. In particular, a request to ‘find the blicket’ could be interpreted very differently by English and Japanese speakers due to differences in the lexical statistics of each language. Although the blicket could technically be either a count or mass noun in English, a child might infer that blicket is likely to be a count noun because count nouns are more frequent than mass nouns in speech (on the assumption that syntactic disambiguation is part of sentence processing). Because count nouns label individuals, the child could then infer that the blicket most likely refers to an individual, and thereby extend the word by shape. By contrast, Japanese children would not be biased by such lexical statistics, because their language does not have count or mass nouns and thus does not require this step of syntactic disambiguation; these children might instead base their word extensions on physical properties of the stimuli alone.

Several studies have provided evidence in support of the lexical statistics hypothesis, by showing that cross-linguistic differences only arise in situations where subjects have to interpret ambiguous words for novel objects: Precisely those situations in which they might draw on lexical statistics. These studies militate against the Whorfian account and suggest that speakers of all languages draw the boundary between individuals and nonindividuals in a similar way. To begin, speakers of English, Japanese, and Mandarin Chinese (another classifier language) provide very similar judgments when asked to rate (on a Likert-like scale) whether stimuli are substances or objects, providing support for the universalist view. Also, in tests of Mandarin-English
bilinguals, the very same subject will extend a novel word more often by shape when tested in English, but by substance when tested in Mandarin. This suggests that language-specific factors like lexical statistics—as opposed to nonlinguistic perception—mediate behavior on word extension tasks.

Perhaps most convincingly, studies have also shown that speakers of languages with mass-count syntax and classifier languages have very similar meanings for known, familiar nouns (e.g., chair, water, and their translation equivalents). While the studies described above have focused on novel words and objects, the Whorfian proposal also clearly predicts that there should be differences in real, existing lexical items between languages, such that count nouns in languages with mass-count syntax should be more likely to individuate than equivalent nouns in classifier languages. To test this idea, Barner and colleagues used a quantity judgment task to test English and Japanese speakers’ interpretations of equivalent, real nouns. Strikingly, subjects of both languages quantified by number for nouns that most often appear in count syntax in English (e.g., key) and by volume for nouns that most often appear in mass syntax (e.g., sugar). For flexible nouns that can appear in both count and mass syntax in English (e.g., stone), English speakers quantified by number when the noun was presented in count syntax (‘Who has more stones?’) and by volume when it was presented in mass syntax (‘Who has more stone?’).

Interestingly, Japanese speakers’ judgments were between the English mass and count judgments, suggesting that these stimuli were ambiguous, and could be construed either as individuals or nonindividuals. Additional studies indicate that, when compared to word extension tasks, quantity judgment tasks are equally sensitive to how speakers of classifier languages and languages with mass-count syntax interpret novel words. Thus, although quantity judgment tasks are capable of detecting cross-linguistic differences, these effects simply are not present for real nouns.

Taken together, these findings suggest that acquiring mass-count syntax does not transform word meanings, such that speakers of languages with mass-count syntax are more likely to have nouns that individuate than speakers of classifier languages. Instead, speakers of different languages draw on a universal set of meanings, with mass-count syntax allowing speakers to select among individuated and nonindividuated meanings, e.g., as in the case of flexible nouns like string or stone. These findings suggest that in languages which lack mass-count syntax, concepts help determine whether a noun will have an individuated meaning, leading speakers of all languages to quantify similarly.

CONCEPTUAL SOURCES OF INDIVIDUATION AND THE ROLE OF PRAGMATIC PROCESSES

Thus far, we have described how syntactic, lexical, and conceptual processes may interact to determine whether a noun labels an individual. In languages like English, mass syntax allows nouns to quantify either by mass/volume (e.g., milk and water) or by number (e.g., furniture), depending on whether these nouns are lexically specified for individuation or not. By contrast, count syntax takes unindividuated lexical items as input, and generates the expectation that they will label individuals (e.g., ‘find a blicket’ implies that blickets are individuals), while lexical concepts provide criteria for determining which individuals the noun labels (e.g., what counts as a blicket). The situation with classifier languages like Japanese—described in the previous section—also underscores the importance of conceptual sources of individuation: Because nouns in these languages never appear in count syntax, they only receive individuated interpretations when such interpretations are provided by concepts. But how do concepts provide criteria for individuation, such that nouns are used to label only kind members, e.g., such that fork labels only whole forks?

Interestingly, several studies suggest that children are surprisingly slow to acquire adult-like criteria for individuating familiar objects. In particular, several studies have found that children are surprisingly willing to use nouns like fork to denote not only whole objects, but also their arbitrary parts. For example, while adults will typically count two whole forks and a third that has been cut into two pieces as either ‘two forks’ or ‘three forks,’ children under age 7 tend to include arbitrary broken pieces in their counts, e.g., resulting in a count of ‘five forks.’ Such behavior occurs even when the broken parts of the object are positioned closely together (making it easy to infer that they form a whole object), when children are asked to count only whole objects, and when they had just counted an object as a single object and it was subsequently broken in front of them. Furthermore, children’s errors extend beyond counting tasks, suggesting that they accept broken pieces as the referents of object nouns more generally. For example, when asked to ‘place a fork in the circle’ or to ‘touch every fork,’ 4-year olds touch arbitrary pieces of forks, and are willing to place a
single fork-piece in a circle.67 Taken together, these studies indicate that unlike adults, children do not restrict object-nouns like fork to label only whole objects.

The findings described above suggest that children may initially assume that all nouns are divisible, and thus resemble count nouns like string and stone: e.g., just as parts of a string can be called strings, so can parts of a fork be called forks. While children could maintain this default assumption for words like string, they could override it for other words, as they acquire full conceptual criteria of individuation for these words (e.g., and learn that parts of a fork do not satisfy the criteria of individuation for fork). However, there are reasons to think that young children’s errors do not reveal a conceptual discontinuity with adults, whereby only adults’ lexical concepts encode full criteria for individuation. In particular, studies have shown that children do not treat all broken parts of an object as members of the object kind: when presented with objects that have been broken into nameable, functional parts, like bicycles broken into frames and wheels, 4-year olds do not count each broken-part as an instance of the object kind.67

The fact that children exclude nameable functional parts from their counts of whole objects suggests a second possibility for how children might overcome a default assumption that all nouns are divisible. Rather than undergoing a conceptual change, children may begin to behave like adults via a developing ability to pragmatically contrast what a speaker said with what they could have said.70–72 For example, when asked to ‘count the bicycles,’ children could infer that wheels should not be counted, because the experimenter would have asked them to ‘count the wheels’ had she wanted them to be counted. Thus, children may succeed at excluding wheels from their counts of bicycles by recognizing that wheel is a better, more informative description of wheels than bicycle. Furthermore, children may not even need to know what functional parts like wheels are called to avoid including them as kind-members: such parts may be conceptualized as distinct kinds of objects because they have their own functions. Consistent with this, children exclude functional object-parts from their counts of whole objects more often than they are able to name them.67

The pragmatic account described above suggests that children should succeed at excluding broken object-parts from their counts of whole objects whenever the object-parts have unique names or functions. To test these predictions, Srinivasan and colleagues73 presented children with sets of whole and broken novel objects, and manipulated both whether the objects were broken into arbitrary or functional parts, and whether the parts were given unique labels or not. As predicted, children were less likely to count broken parts as instances of the whole object—and were thus more adult-like—when the parts were functional or had their own labels, suggesting that labels and functional information are each sufficient for excluding parts.

The findings of Srinivasan and colleagues73 suggest that children may initially include arbitrary parts, like pieces of fork, in their counts of whole objects, both because such parts are not functional, and because children cannot access better labels for these parts than whole object labels. To indicate arbitrary object-parts, adults typically use measure phrases, like piece of a fork, or half of a fork, but young children may not have acquired such phrases or recognized that these phrases are more informative descriptions of arbitrary parts than whole object labels. Once children learn measure phrases, they could infer that a request to ‘count the forks’ implies that arbitrary fork-pieces should be excluded, because the request would have otherwise used a measure phrase.

However, subsequent studies have shown that 4-year olds understand measure phrases but still make counting errors,73 suggesting that their errors do not arise from not having learned measure phrases. For example, when presented with a set (e.g., of two whole shoes or of a single shoe cut in half) and a pair of descriptions (e.g., ‘Farmer Brown says it’s two shoes’ vs ‘Captain Blue says it’s two pieces of shoe’), 4-year olds choose the more informative description. This suggests that children may count arbitrary parts as instances of whole objects not because they do not understand measure phrases, but instead because they do not spontaneously generate measure phrases as alternative descriptions of arbitrary parts, relative to whole object labels (for a related idea from the case study of Gricean scalar implicature, see Ref 74).

To test whether children’s errors arise from failing to access measure phrases while counting, Srinivasan and colleagues73 explored whether children exclude pieces more often from their counts of whole objects when descriptions of objects and pieces have first been primed. In the critical priming condition, children were first shown a whole object (e.g., a fork) and an arbitrary part of that object (a piece of a fork), and were asked to indicate which was the referent of a measure phrase (‘Can you point to the piece of a fork?’) and whole object label (‘Can you point to the
fork?’). Then, they were asked to enumerate a set using the whole object label (‘Can you count the forks?’), which contained two whole objects and a third that had been broken into two or three pieces. Consistent with the idea that children’s counting errors stem from difficulty with accessing alternative descriptions, the priming manipulation made children’s counting significantly more adult-like.

These findings suggest that children do not begin to behave like adults—and restrict object–count nouns to label only whole objects—simply by virtue of a conceptual insight. Critically, children who received the priming manipulation were not told which of the items were pieces and which were wholes, making it unlikely that their concepts changed during the task. Instead, children may overcome a default assumption that nouns are divisible not solely by acquiring full criteria for individuation, but also by spontaneously contrasting whole object labels against other alternative descriptions, including labels for functional parts (e.g., wheel), and measure phrases for arbitrary parts (e.g., piece of fork, half of shoe, etc.). This approach to how lexical concepts restrict reference to specific individuals is unique in that it lightens the explanatory burden traditionally placed upon concepts\textsuperscript{25}: rather than providing full criteria for individuation, concepts may only need to provide partial conditions that are enriched pragmatically.

CONCLUSION

Guided by the idea that linguistic structure can provide a window onto how we conceptualize the world, scholars have long theorized about the relationship between mass–count syntax and how we individuate entities in the world.\textsuperscript{1–14,28,33,34,44,57} In this study, we have argued that a full understanding of how language encodes individuals requires considering how syntactic, lexical, conceptual, and pragmatic processes interact.

We began by arguing that acquiring mass–count syntax does not allow children to make an ontological distinction between objects and substances. Furthermore, a distinction between objects and substances—or a more abstract one between individuals and nonindividuals—does not form the semantic basis of the mass–count distinction. Instead, studies of quantity judgment and word extension suggest that the mass–count distinction is asymmetric, such that while count nouns denote individuals, mass nouns can denote both individuated and nonindividuated entities.\textsuperscript{17–19,25,42} We proposed that these generalizations are best understood in terms of a model in which syntax and lexical roots interact. Count syntax takes unindividuated lexical representations as an input and specifies quantification over individuals, relying on concepts to specify what these individuals are. On the other hand, mass syntax leaves the measuring dimension to individual lexical roots, such that some nouns like water quantify according to mass/volume, while others like furniture are prespecified for individuation and quantify according to number.

Returning to the Whorfian proposal, we also explored whether acquiring mass–count syntax shapes whether entities are construed as individuals, by considering data from languages that lack mass–count syntax. Here, we found that speakers of languages with mass–count syntax are not more likely to have nouns that individuate than speakers of classifier languages,\textsuperscript{16,21} and that previous Whorfian effects\textsuperscript{2,4,58} stem from the different lexical statistics of these languages.\textsuperscript{16,21,63} These findings suggest that in languages that lack mass–count syntax, concepts help determine whether a noun will have an individuated meaning, leading speakers of all languages to quantify similarly. The power of mass–count syntax, we argued, is that it allows speakers to select among individuated and nonindividuated meanings, as in the case of flexible nouns like stone. Finally, we explored how concepts provide criteria for determining which individuals are in a noun’s denotation, given that count syntax may only generate the expectation that a noun will label individuals, without specifying what those individuals are. Drawing on data from children’s early quantification of objects\textsuperscript{64–67,73} we argued that lexical concepts do not provide necessary and sufficient conditions for reference, but instead encode partial criteria that are filled out pragmatically, by contrasting alternative descriptions of objects and their parts.

In sum, research into the mass–count distinction provides one of the clearest windows to date into how different domains of language interact to encode quantity. The progress made on this topic has continued in recent years, as scholars document how quantity is encoded across previously unexplored languages.\textsuperscript{20,22,26,76,77} This work has shown that languages can encode a mass–count distinction in different ways,\textsuperscript{20,26,44,77} and that some languages may fail to encode one altogether.\textsuperscript{22} These findings raise new and exciting questions about how the mass–count distinction is encoded in language, and promise to open additional avenues in our understanding of how language encodes reality.
NOTE

a Some researchers have noted that NP quantification is also affected by other factors, like object function, and that these factors may have especially strong effects on mass nouns, \(^4\,5,6,8\), while others have noted that such accounts could not explain differences between mass and count nouns. \(^7\).

REFERENCES


77. de Oliveira RP, Rothstein S. Bare singular noun phrases are mass in Brazilian Portuguese. *Lingua* 2011, 121:2153–2175.