The Time-Course and Cost of Telicity Inferences

Andrea S. Proctor (a-proctor@northwestern.edu)
Department of Psychology, Northwestern University
2029 Sheridan Road, Evanston, IL 60208 USA

Michael Walsh Dickey (m-dickey@northwestern.edu)
Department of Communication Sciences and Disorders, Northwestern University
2240 Campus Drive, Evanston, IL 60208 USA

Lance J. Rips (rips@northwestern.edu)
Department of Psychology, Northwestern University
2029 Sheridan Road, Evanston, IL 60208 USA

Abstract
Recent evidence suggests that perceivers have consistent intuitions regarding the boundedness properties of objects and events (Solomon, Proctor, & Rips, in preparation). This paper presents a self-paced reading study examining the speed and accuracy with which readers draw such telicity inferences during on-line language comprehension. Participants read sentences containing either a consumption verb (“consume”) or an observation verb (“monitor”) followed by either a mass or a count object (“ice water” vs. “ice cube”). Each sentence ended with an adverbial phrase that was either consistent or inconsistent with the telicity of the preceding event description (“in” or “for” adverbials), along with a comprehension question. Reading-time results suggest that comprehenders are slow to draw telicity inferences, even when the type of verb unambiguously determines the telicity of the sentence. However, responses to post-sentential comprehension questions suggest that verb and noun information together have a surprisingly robust influence on comprehenders’ telicity inferences, even in the face of supposedly unambiguous adverbial information. Together, these results suggest that comprehenders make use of all relevant information in making telicity inferences, but that they do so much more slowly than strongly incremental models of natural language understanding would predict (e.g., Marslen-Wilson & Tyler, 1980).

Introduction
We readily distinguish two types of physical entities in the world—individuated objects and substances. We refer to these two types of entities with different types of nouns: substances are often referred to with mass nouns (e.g., tea) and objects with count nouns (e.g., cat). These distinctions seem to rest on the boundaries of the physical entities: count nouns typically refer to objects with well-defined boundaries, such as mouse or iceberg, while mass nouns typically refer to substances1 without clear boundaries, such as mud or water.

1 It should be noted, however, that this distinction between mass and count nouns does not strictly coincide with the substance/object distinction. Jackendoff (1991) notes that, in addition to substances, aggregates of individuated objects can act like mass nouns. For instance, the terms cattle and change behave like mass terms, even though each refers to individuated objects (cows and coins, respectively) and not to an unindividuated substance.

The domain of events can be divided up similarly. Events can be classified according to whether or not they have an endpoint, or temporal boundary. Actions described by atelic verbs have no inherent endpoint or boundary; these actions have the potential to go on without end (e.g., singing). Verb phrases describing atelic or unbounded events go naturally with for adverbials, which describe the duration of an event, and less naturally with in adverbials, which presuppose the endpoint of an event. “She sang for an hour” sounds much more natural than “She sang in an hour.” Actions described by telic verbs such as delivering, on the other hand, have an inherent endpoint; once an object has arrived at its destination, delivering has reached its end, and cannot logically continue. Verb phrases describing telic or bounded events go naturally with in adverbials and less naturally with for adverbials: “She delivered the package in an hour” sounds much more natural than “She delivered the package for an hour.” These distinctions among verbs or verb phrases are commonly referred to as lexical aspect (e.g., Vendler, 1967; Dowty, 1979).

The current experiment explores how these lexical aspect distinctions are computed during sentence comprehension. It also examines when perceivers draw inferences about the boundedness of events, just as they must draw inferences about the boundaries of physical entities (see Solomon, Proctor, & Rips, in preparation). There has been some previous work exploring the cost of modifying or retracting such inferences once they have been drawn: Piñango, Zurif, and Jackendoff (1999) and Todorova, Straub, Badecker, and Frank (2000) demonstrate that encountering information (such as a for adverbial) that forces an event to be construed as atelic causes processing difficulty if previous information had suggested it was telic. The current experiment uses this effect to explore when telicity inferences are drawn on-line.
**Parallels Between Noun and Verb Contrasts**

Several authors have noted that strong parallels exist between mass nouns and atelic events, and between count nouns and telic events (Bach, 1986; Langacker, 1987; Vendler, 1967). Just as masses have no intrinsic physical bound, atelic events, such as running, painting, or watching, have no intrinsic temporal bound. In contrast, just as counted objects have inherent physical boundaries, telic events, such as delivering, drowning, or walking a mile, all have an intrinsic temporal bound. One cannot continue delivering after a package has reached its destination, drowning after one is dead, or walking a mile after that distance has been crossed. Once the endpoint is reached, the action is completed. The present experiment is concerned with one subtype each of atelic and telic verbs: activities and accomplishments, respectively.

Another key parallel between actions and objects involves the extent to which a part of an object or action can be considered to be in the same category as its whole. For both masses and activities, a subpart (down to some lower limit) of the whole is qualitatively equivalent to the whole—any part of chocolate sauce is still chocolate sauce, just as any part of eating is eating. This subpart or subinterval property (Bennett & Partee, 1978) does not hold for counted objects and accomplishments, however—any part of an aluminum boat is not, itself, an aluminum boat, nor is any part of lighting a fire (e.g., crumpling up newspaper) itself lighting a fire.

**Interactions Between Noun and Verb Boundaries**

The physical boundaries of objects influence the temporal boundaries of events affecting them. Several authors have noted that the telicity or boundedness of an event often depends on whether the verb describing it takes a mass or count noun as its object (Pustejovsky, 1991, 1995; Verkuyl, 1993). When a consumption verb, such as eat, takes a count noun as its object, readers should infer that the VP is an accomplishment—the depletion of the object must end when the object’s boundary is reached. When such a verb takes a mass noun as its object, however, the VP is an activity—since the substance is unbounded, the depletion could potentially go on indefinitely. These telicity shifts only hold for a subset of verbs such as consumption and creation verbs, which describe events that cannot easily be repeated and that entail an irreversible effect on their objects (Krifka, 1998). Verbs that do not entail an irreversible effect on their objects should not demonstrate such an aspectual shift. For example, whether a mass or count noun appears as the object of a verb of observation (e.g., watching) should not have an effect on the VP’s telicity—the action of watching a mug should be just as unbounded as the action of watching soup.

To date, there is limited evidence regarding how perceivers draw these telicity inferences during comprehension. Solomon and her colleagues (in preparation) provide evidence from off-line reasoning tasks suggesting that readers are sensitive to the boundedness of different events, and that their reasoning about event boundedness parallels their reasoning about the boundedness of physical objects. Further, readers make inferences about lexical aspect on-line, re-interpreting their default assumptions regarding the boundedness of an event in order to bring it in line with the temporal characteristics of the context in which it appears (Pinango et al., 1999; Todorova et al., 2000). For example, “jump” might be interpreted to be an iterative action when it appears in the sentence “He jumped all day.”

Studies examining aspectual coercion have demonstrated that sentential aspect is sensitive to parts of the sentence other than the verb. However, they do not directly examine how the boundedness of a verb’s object or the verb itself is capable of influencing telicity inferences. If perceivers draw inferences about a sentence’s temporal profile incrementally, they should show early sensitivity to the difference between observation verbs (whose boundedness does not depend on the properties of the following object) and consumption verbs (whose boundedness does depend on the object). Similarly, if perceivers are actively and predictively computing telicity based on verb and object information, they should show evidence of an interaction of verb and object information at the position of the object: perceivers’ comprehension of a sentence should be affected by whether an object is a mass or count noun only if the verb preceding it is a consumption verb, not if it is an observation verb. Strongly incremental views of language understanding (e.g., Marslen-Wilson & Tyler, 1980; Clark, 1996) predict such early inferencing.

**Overview**

Is it the case that when we read a sentence in which an (unbounded) activity verb, describing the consumption of its object, takes a (bounded) count noun as its object, we then interpret the sentence as if the verb phrase were bounded? If, for instance, we read “Carol ingested Henderson Foods’ rice cake merrily for ten minutes” will we interpret that action as being more temporally bounded than we would if we had read that she had ingested Henderson Foods’ rice cereal? When do we make these inferences? Do we begin to draw inferences regarding telicity as soon as we encounter a verb? Do we compute a sentence’s aspect as soon as we have both the verb and noun information? Or do we hold off making inferences until all potentially informative information is available, until late in the sentence? Previous results suggest that drawing telicity inferences can be cognitively costly (see Pinango et al., 1999; Todorova et al., 2000); the present experiment uses this finding to address these questions.

**The Experiment**

We presented participants with sentences describing characters either consuming or observing a mass or a counted object (see Table 1 for examples). Sentences were divided into five segments, and participants read through
Table 1: Sample Set of Sentences

<table>
<thead>
<tr>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Segment 3</th>
<th>Segment 4</th>
<th>Segment 5</th>
<th>Adverb</th>
<th>Verb</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Leslie consumed</td>
<td>Polar Purity’s</td>
<td>ice water</td>
<td>with zeal</td>
<td>for eight minutes.</td>
<td>For</td>
<td>Telic</td>
<td>Mass</td>
</tr>
<tr>
<td>b Leslie consumed</td>
<td>Polar Purity’s</td>
<td>ice cube</td>
<td>with zeal</td>
<td>for eight minutes.</td>
<td>For</td>
<td>Telic</td>
<td>Count</td>
</tr>
<tr>
<td>c Leslie monitored</td>
<td>Polar Purity’s</td>
<td>ice water</td>
<td>with zeal</td>
<td>for eight minutes.</td>
<td>For</td>
<td>Atelic</td>
<td>Mass</td>
</tr>
<tr>
<td>d Leslie monitored</td>
<td>Polar Purity’s</td>
<td>ice cube</td>
<td>with zeal</td>
<td>for eight minutes.</td>
<td>For</td>
<td>Atelic</td>
<td>Count</td>
</tr>
<tr>
<td>e Leslie consumed</td>
<td>Polar Purity’s</td>
<td>ice water</td>
<td>with zeal</td>
<td>in eight minutes.</td>
<td>In</td>
<td>Telic</td>
<td>Mass</td>
</tr>
<tr>
<td>f Leslie consumed</td>
<td>Polar Purity’s</td>
<td>ice cube</td>
<td>with zeal</td>
<td>in eight minutes.</td>
<td>In</td>
<td>Telic</td>
<td>Count</td>
</tr>
<tr>
<td>g Leslie monitored</td>
<td>Polar Purity’s</td>
<td>ice water</td>
<td>with zeal</td>
<td>in eight minutes.</td>
<td>In</td>
<td>Atelic</td>
<td>Mass</td>
</tr>
<tr>
<td>h Leslie monitored</td>
<td>Polar Purity’s</td>
<td>ice cube</td>
<td>with zeal</td>
<td>in eight minutes.</td>
<td>In</td>
<td>Atelic</td>
<td>Count</td>
</tr>
</tbody>
</table>

In the absence of slowing at segments three or four, we could still determine that inferences about aspect were being made based on a combination of a consumption verb with a mass or count noun if we were to find slowing due to mismatching of grammatical information at segment five. If participants are making inferences about the aspect of the sentences (e.g., inferring that eating a chocolate bar is an accomplishment) and are then presented with a final adverb (e.g., for ten minutes) that contradicts this inference, their reading times at that final segment may be slowed.

As an additional test of readers’ inferences about the aspect, we presented participants with a follow-up question immediately after they had read through each sentence. For example, after reading about Leslie consuming ice water, participants were presented with the question, “After four of those eight minutes, had Leslie actually ingested Polar Purity’s ice water?” The questions were modeled after an inferential test (Dowty, 1979) that distinguishes between activities and accomplishments by assessing whether the subinterval property applies to the action. If a participant believes that the action was an activity, he or she should be willing to ascribe the subinterval property to it. We thus expected that participants would be more likely to respond “yes” to the follow-up question after they had read about Leslie consuming ice water than if they had read about Leslie consuming Polar Purity’s ice cube, since the consumption of a bounded object should lead participants to interpret the action as telic. Observation verb sentences should have the subinterval property ascribed to them regardless of whether they contain a mass or count noun, as the action has no effect on the object. We expected the pattern for sentences ending with the in-adverb to be qualitatively similar to that for sentences ending with the for-adverb. However, overall likelihood of an atelic response should be lower for in, given the strong association of in-adverbials with telic actions.

**Methods**

**Procedure** In this study, participants completed a self-paced reading task. They read through sentences, presented on a computer screen while we recorded their reading times. We instructed participants to read through each segment at their normal pace, and to progress through the segments by pressing the spacebar. The segments appeared sequentially...
and disappeared from view once they had been read; participants could not return to a previously-viewed segment.

A follow-up question appeared immediately after participants had read the final segment of the sentence. The follow-up question appeared with two possible responses (“yes” and “no” for the experimental sentences), one appearing on the left, and the other on the right. For each list, participants saw “yes” on the left for half of the experimental sentences, and “no” on the left for the other half. We instructed participants to press one key (d) if they felt that the response on the left was correct and another (k) if they felt that response on the right was the better choice.

Each participant saw the sentences and their associated follow-up questions in a different random order.

Materials We constructed 80 sentences, all describing a character performing some action on an object or substance over a specified interval. For each of these sentences, we varied the type of verb (observation verbs vs. consumption verb), the boundedness of the noun (mass vs. count noun), and the final adverb (for X minutes vs. in X minutes), yielding a set of eight variations on each of the 80 base sentences. Each sentence contained five segments (an example set of segmented sentences appears in Table 1). The 640 experimental sentences were separated into eight lists, such that each of the eight variations of any given base sentence was assigned to a different list, and each list contained ten of each of the variation types. Thus, participants saw only one version of each base sentence, but saw an equal number of each of the verb/noun/adverb combinations. Each list also contained 32 additional sentences, unrelated to the experimental sentences, which served as filler items (e.g., “At the break, | Emily | had already | finished the memo, | to her boss’s relief.”)

Follow-up questions were constructed for all the sentences. The follow-ups asked whether the character had actually completed the specified action halfway through the mentioned interval (an example set of follow-up questions is presented in Table 2). If, during the course of reading a sentence, the participant inferred that the described event was atelic, then the subinterval property should apply, and the expected response would be “yes”. If, on the other hand, the participant inferred that the described event was telic, he or she should respond “no”.

| a, c | After four of those eight minutes, had Leslie actually consumed Polar Purity’s ice water? |
| b, f | After four of those eight minutes, had Leslie actually consumed Polar Purity’s ice cube? |
| c, g | After four of those eight minutes, had Leslie actually monitored Polar Purity’s ice water? |
| d, h | After four of those eight minutes, had Leslie actually monitored Polar Purity’s ice cube? |

Participants Forty-eight undergraduate students enrolled at Northwestern University participated in this experiment. Participation was part of a course requirement in an introductory psychology course. All participants were native English speakers.

Results

An examination of the reading time data suggests that readers did not begin generating telicity inferences as soon as they encountered the relevant verb or verb+object information. Reading times were, however, longer for sentences in which the final adverbials were inconsistent with the telicity of the preceding verb+object combination, indicating that the verb+object information was used in generating inferences about the telicity of these described events. Responses to follow-up questions suggest that these inferences were surprisingly robust: even in cases where the sentence-final adverbial conflicted with the telicity of the preceding verb phrase, participants showed some evidence of having stuck with their original telicity inference.

Analyses on reading times and responses to follow-up questions were computed separately using participants and items as random factors.

Reading-Time Analyses Mean reading times for each segment are presented in Figure 1. Reading times that were over 10 seconds or under 100 milliseconds were excluded from the analysis. These responses consisted of less than 2% of the data.

Segment-by-segment reading time analyses revealed a significant effect of verb type at segment one (F(1,47) = 7.21; F(1,79) = 4.11, p < .05 for both); however, this effect was in the direction opposite to that expected under a strongly incremental view—observation verbs were read more quickly than consumption verbs. There was no evidence of additional inferencing work going on in the atelic conditions.2

Also speaking against the predictions of an incremental account is the finding that there were no significant differences at segment three (the count/mass noun segment) or segment four (the manner adverbial segment), indicating there is no immediate processing cost associated with drawing telicity inferences based on the integration of noun and verb information.

Despite the lack of immediate processing, an examination of the reading times for the final segment (see Figure 2) provides some evidence that participants were combining the verb and noun information to generate inferences about the telicity of the events. Analyses of segment five reading times reveal evidence of processing costs when participants encounter grammatical information conflicting with telicity inferences. A 2x2x2 repeated measures

2 This main effect may have been the result of a confound with word frequency—the mean frequency (Kucera & Francis, 1967) of the observation verbs was significantly higher than that of the consumption verbs (t(59) = 3.64, p < .01).
ANOVA revealed a three-way interaction between noun, verb and final adverb (\(F_{(1,47)} = 4.29, F_{(1,79)} = 4.46, \text{both } p < .05\)). This interaction reflects both the relatively long segment five reading times for sentences whose final adverbs conflict with the telicity information contained in earlier segments (types e, g, and h). This difference was confirmed with planned contrasts (\(F_{(1,47)} = 11.23, F_{(1,79)} = 33.01, \text{both } p < .01\) for both).

Looking at the for-adverbial conditions alone, there was a two-way interaction of verb and noun type with slower reading times for sentences with consumption verbs and count nouns than sentences with consumption verbs and mass nouns, though it was only marginally significant in the items analysis \(F_{(1,47)} = 9.72, p < .01; F_{(1,79)} = 3.71, p < .06\). This suggests that participants had successfully drawn telicity inferences based on verb and noun information earlier in the sentence.

Follow-up Question Data

Responses to the follow-up questions indicate that the participants were sensitive to the experimental manipulations. Repeated measures ANOVAs revealed main effects for final adverb \(F_{(1,47)} = 66.84; F_{(1,79)} = 757.12, \text{both } p < .001\), verb \(F_{(1,47)} = 64.08; F_{(1,79)} = 131.67, \text{both } p < .001\) and, though only marginal in the items analysis, for noun \(F_{(1,47)} = 4.27, p < .05; F_{(1,79)} = 3.14, p = .08\). This pattern of responses indicated that, as expected, participants were more likely to respond “yes” (indicating an atelic interpretation) to the questions following sentences that contained for adverbs, atelic verbs, and mass nouns than they were for sentences containing in adverbs, telic verbs, and count nouns, respectively (see Figure 3). There were no significant interactions. Planned t-tests comparing responses to questions following mass and count versions of the critical consumption/or sentences revealed that participants were more likely to attribute the subinterval property to an event if it involved the consumption of a mass than a counted object, though the difference was only marginal in the items analysis \(t_{(47)} = 2.23, p < .05; t_{(79)} = 1.97, p = .05\), indicating that they were more likely to treat a consumption/mass action as an activity than a consumption/count action. This pattern occurred despite the presence of the for-adverbial, which should force an atelic interpretation.

General Discussion

An incremental view of language understanding would predict that participants should begin drawing telicity inferences as soon as those inferences are licensed by the text. In the current study, such a view predicts that participants should show slowed reading times at the first segment (the verb segment) for observation verbs (relative to consumption verbs) since such verbs license immediate inferences about the telicity of the events they describe, whereas consumption verbs do not. The present study found little support for this prediction—the main effect of verb found at the first segment was in the opposite direction of that predicted.
An incremental view further predicts an interaction between noun and verb at the third and fourth segments (the noun and manner adverbial segments). If drawing inferences about the telicity of a VP based on verb and noun boundaries were costly, we would expect to see a slowing in reading times as soon as conflicting information appeared (e.g., when a count noun followed a consumption verb). This prediction also failed—there was no such interaction. Nevertheless, participants were using the combined verb and object information to make inferences about the boundedness of events. Both reading time differences at segment five and responses to follow-up questions provide support for these inferences. In the former case, reading times increased when a final adverbial was inconsistent with the verb+object combination. In the latter case, participants were more likely to agree with the subinterval property for the consumption of a mass than for the consumption of a counted object.

Conclusions

Earlier work shows that there is a processing cost for drawing inferences about the telicity of events (Piñango et al., 1999; Todorova et al., 2000). The present study investigated the time course of such inference-drawing. Two strong possibilities presented themselves at the outset: participants could either make inferences early, as the relevant information was presented to them (as in an incremental account) or, alternatively, they could hold off making telicity inferences until late in sentence processing, when all information was available (minimally, until after verb and noun information was available; Pustejovsky, 1991, 1995; Verkuyl, 1993). The reading time results from the present study support the second alternative—there is no indication that participants made rapid use of either the verb or verb+object information to draw boundedness inferences. Instead, it seems that all the costly inferencing work was carried out at the final segment. This finding is consistent with other work: for example, Todorova and her colleagues (2000) find no cost for combining a telic verb (such as send) with a bare plural noun (such as letters), even though the bare plural forces an atelic interpretation for the verb phrase. The absence of such a cost is surprising under strongly incremental views of natural language interpretation (Marslen-Wilson & Tyler, 1980; Altmann & Kamide, 1999). This pattern is also consistent with the possibility that drawing (or delaying) inferences regarding telicity may be a relatively cost-free process, much as delaying choosing among different metonymic or metaphoric uses of a polysemous noun (such as newspaper) (Rayner & Frazier, 1989). In the domain of telicity, participants are willing to hold off on doing costly inferences until they are forced to make an interpretation.

Acknowledgments

This work was supported by NSF grant SES-9907414. We would like to thank Sam Day for his assistance with programming, and three anonymous reviewers for detailed and useful comments on the paper.

References


