Chapter 11

Is the weak definite a generic? An experimental investigation

Thaís Maíra Machado de Sá
Universidade Federal de Minas Gerais

Greg N. Carlson
University of Rochester

Maria Luiza Cunha Lima
Universidade Federal de Minas Gerais

Michael K. Tanenhaus
University of Rochester
Nanjing Normal University

We discuss the properties of WEAK DEFINITE noun phrases, definite noun phrases (henceforth DP) which do not uniquely refer to an individual referent. Since one of the properties of generic noun phrases is that they do not uniquely refer, we asked whether weak definites might in fact be a form of generic noun phrase. We adopted a quantitative and experimental approach conducting a corpus analysis and four experiments that were designed to assess whether weak definites differ from DPs that are generic, weak and regular definites. A corpus analysis by de Sá et al. (2016) showed that generic DPs and weak definites are not in complementary distribution. A follow-up analysis on verb *aktionsart* showed that most weak definites appear in telic or activity DPs. The experiments also compared matched sentences with weak, regular and generic reading DPs. These studies do not find similarities between weak definites and generics. We conclude that weak definite noun phrases are not generics.

1 Introduction

Definite reference has played a central role in linguistics, the philosophy of language and in psycholinguistics (Russell 1905; Strawson 1950; Donnellan 1966; Clark & Marshall 1981; Heim 1982; Aguilar-Guevara & Zwarts 2013). Modulo some nuanced differences in the treatment of definite reference, there is general agreement that definite noun phrases carry a "familiarity", "uniqueness" or "identifiability" condition; the referent of a definite referring expression should be uniquely identifiable within a referential domain. In Example (1), *the hospital* denotes only one hospital in the world, being *unique*, and it is known by the interlocutors, being *familiar*.

(1)   *Workers picketed **the hospital** to protest layoffs.*

However, so-called *weak definite*[1] noun phrases (Carlson & Sussman 2005) such as *the hospital* in (2) violate uniqueness: the speaker does not need to have any specific hospital in mind when she utters *the hospital*. Moreover, John and Bill could even be going to different hospitals.

(2)   *John went to **the hospital** and so did Bill.*

It is also known that reference in definite noun phrases can be generic. In those cases, the definite noun has uniqueness of a kind, i.e. it denotes a kind, not an individual referent. *The hospital* in (3) is an example, because it does not have an unique individual referent, but a kind referent, *the hospital* is a kind of place.

(3)   *In the XVIII century, hygiene rules were introduced into **the hospital** in the Western world.*

For Aguilar-Guevara & Zwarts (2011: 193) weak and generic definites would be "different faces of a same phenomenon", because both of them would have the uniqueness of a kind property, denoting a kind. Indeed, if the lack of individual reference in weak definites can be reduced to the fact they are generic definites, it would be the most straightforward means of accounting for this lack of individual reference.

---

[1]Poesio (1994) was the first to use the name *weak definites*, questioning the Russellian uniqueness (1905) and Heim's familiarity (1982). He noted that in sentences like *John got these data from the student of a linguist* there is no need to have familiarity or characterize a single individual to *the student* in order to understand the sentence. He named this class of definites *weak definites*. Carlson & Sussman (2005) adopted the *weak definites* term, observing that weak definites lack uniqueness.

The current work does not directly address the specific analysis proposed by Aguilar-Guevara & Zwarts (2011). Instead we address the basic question of to what extent weak definites share the properties of generic noun phrases and regular noun phrases.

In this chapter, we employ empirical means to evaluate the hypothesis that definite generics and weak definites are the same phenomenon. We will examine corpus data form Brazilian Portuguese, and experimental data from English to evaluate this question.

We begin with a brief summary of the properties of weak definites.

2 Weak definites

The term *weak definite noun phrase* is used here to describe a certain kind of construction that Carlson and collaborators (Carlson & Sussman 2005; Carlson et al. 2006; 2013 and Klein et al. 2013) have been working on for some time under this designation. The contrasting class of definite noun phrases is called *regular definites* (sometimes "strong definites"), meaning that they trigger the familiarity/uniqueness presuppositions commonly focused in the literature on definite descriptions. The term *weak definite noun phrase(s)* is often elided to simply *weak definite(s)*, but we wish to be clear that we do not use this term in the present context to refer to just any noun phrase which, in a language differentiating "strong" vs. "weak" definite article forms, has the definite article in the "weak" form. When we wish to refer to the morphological forms of definite articles, we will do so explicitly.

Besides failing to trigger uniqueness presuppositions, these noun phrases, among other properties, must occur in construction with a specific verb or preposition, may only occur in the singular form or the plural form but not both, and are not subject to restrictive modification.[2] They appear to have the semantic truth-conditions of narrow-scope indefinites, and normally trigger semantically "enriching" implications – i.e. there is a non-compositional aspect to their meaning. Finally, the constructions appear to have a more "eventive" meaning than the corresponding compositional constructions, a matter we try to pin down a bit more precisely below.[3]

Our work was motivated in part by the incorporation hypotheses proposed by Carlson and colleagues. Weak definite noun phrases are treated as an incorpo-

---

[2]See Aguilar-Guevara (2014) for insight into the allowable modifiers.

[3]The constructions under consideration have a number of characteristics that are summarized in Carlson et al. (2006).

rated structure by Carlson et al. (2013) and Klein et al. (2013), in which the noun phrase and the verb have the semantics of an incorporated event in which the article, definite or indefinite, takes scope over the incorporated structure. This analysis unifies the observation that weak definites need not uniquely refer and the observation that they evoke habitual events associated with the noun. It also provides an explanation for the role of the definite article and makes the novel prediction that the same noun phrases that can have a weak definite interpretation can also appear in "weak indefinite" structures, which are incorporated structures that have properties more characteristic of an indefinite than a definite. Crucially this approach assumes that weak definites do not have the same properties as generic DP.

In an attempt to better understand the role of the definite article in the determined phrase and in the incorporated construction, we conducted a corpus analysis and a set of experiments that examined whether weak definites exhibit properties of generics (§3). Then, we report the results of four experiments (§4).

3 Corpus analysis

In order to observe if weak definites would pattern with generic definites, de Sá et al. (2016) analyzed data on a Brazilian Portuguese (BP) corpus. Four-hundred occurrences of 31 words, which may present the weak reading in BP (e.g. *the hospital*), were analyzed. They analyzed whether the word was determined by a definite article, and if so, whether the DP reading was weak (Carlson & Sussman 2005), strong – or regular – (Russell 1905), or generic (Carlson 2006). They then looked at the distribution of those three kinds of definites. As expected, the regular reading is significantly more frequent than the others, 45.6%, but surprisingly, according to the categorization criteria, the weak DPs occur significantly more often than the generic ones, 33.7% versus 27.5%.

The authors also described the DP's syntactic function – subject, object, adjunct – for occurrences of weak, regular and generic definites in the corpus analysis. The goal was to compare the distributional properties of weak definites, generic DPs and regular definites. They evaluated two hypotheses:

1. If weak definites are in fact generics, then generic DPs and weak definites should either occur in the same environments or be in complementary distribution with one another, indicating that they are variations of the same linguistic type.

2. The second hypothesis was motivated by an analysis that weak definites undergo semantic incorporation proposed by Carlson et al. (2013). The semantic incorporation hypothesis predicts that weak definites should occur primarily as the object of a verb or a preposition but rarely should occur in subject position.

They found that generics (Figure 1A) are more uniformly distributed between subject (25.1%) and object (20.3%), being adjuncts most frequently (54.6%). Regular definites showed the same overall pattern (Figure 1B), presenting a significant majority of adjuncts (43.7%), followed by objects (31.3%), and subjects (25%). Weak definites presented a different distribution in which they appear as adjuncts (45.7%) as often as objects (46.6%). Weak definites, however, seldom appear as subjects. Only 7.2% of the occurrences were as subjects, significantly less than the other categories (Figure 1C).
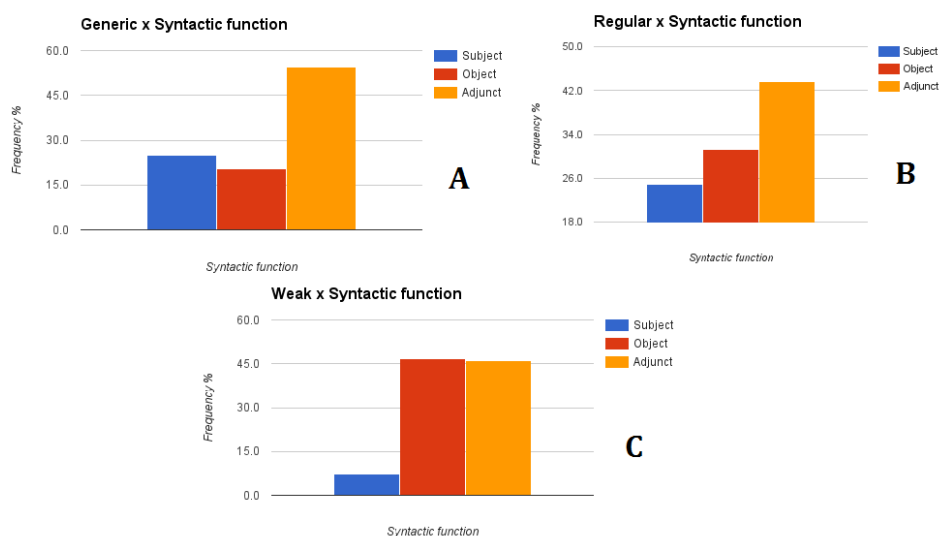

Figure 1: Definite types and syntatic function – Generic definites (A), Regular definites (B) and Weak definites (C) (de Sá et al. 2016: 114, 115)

The authors argued that the weak definites' high occurrence in adjunct and in object position could be interpreted as a reflex of an incorporation process, as proposed by Carlson et al. (2013) and Klein et al. (2013). But the fact that this kind of definite could also be found in subject position is a problem for the incorporation analysis. The data also did not point to a complementary distribution

between weak and generic definites, which could be argued to provide support for the claim that they are the same phenomena.

3.1 Aktionsarten analysis

As a following analysis to the syntactic function analysis made by de Sá et al. (2016), we, with the same tagged corpus, used the verb to analyze the semantics of the clause in which the definite noun occurred. The verb aktionsarten[4] was the semantic property we focused on motivated by the incorporation analysis, which claims that weak definites are incorporated in event or activity verbs (Carlson et al. 2013).

Our hypothesis was that in aktionsart analyses, the semantic incorporation hypothesis predicts that weak definites (but not generic DPs) should primarily occur with activity and telic verbs, but not with state verbs. We also compared weak definites with generics, which are usually found in clauses with state verbs (Carlson 2006), to see if there is a complementary distribution between those categories.

For the same 2196 occurrences (of 31 words which could have generic, weak and regular readings)[5] from de Sá et al. (2016) we analyzed the lexical aspect of the verb for the clauses containing the definite expression.

The verbs were classified as *state*, *activity*, or *telic* (achievement and accomplishment), based on Vendler (1957). We classified as *state* verbs those that do not denote an action, for example the verb *ter* in BP, in the Example (4):[6] *tem* does not have a process which unfolds during time, it does not denote action and if we consider its thematic role, then the subject, *the school* is not an agent.

(4) Brazilian Portuguese

*Além   do   atendimento pedagógico, a   escola **tem** responsabilidades*
Beyond of+the service   pedagogical the school **has** responsibilities
sociais.
social

'The school has social responsibilities, which goes beyond the pedagogical service.'

---

[4]We analyzed Vendler (1957) aktionsarten's categories: state, activity and telic (achievement and accomplishment).

[5]Extracted from the *ptTenTen* corpus, in the platform *Sketch Engine*. See more information in de Sá et al. (2016).

[6]From here until the end of this section all the examples are from our data.

The *activity* verbs are actions which do not need a conclusion point, as the verb *nadar* in Example (5): *nadavam* is an action that unfolds during time, but it does not have a finishing point.

(5) Brazilian Portuguese
*Os alunos **nadavam** todo dia na   escola.*
The students **swam**   every day in+the school
'The students swam every day in the school.'

We classified as *telic* the action verbs that needed a finishing point, as *quebrar*, in Example (6): *quebraram* is an action that requires a conclusion point.

(6) Brazilian Portuguese
*Os vândalos **quebraram** a escola durante a festa.*
The vandals **broke**   the school during the party
'Vandals broke the school during the party.'

In addition to the notion of aktionsart proposed by Vendler (1957), we used the aspectual tests in Dowty (1979) to distinguish one category from another in our analysis. As the Dowty tests are proposed for English, we used a version proposed by Wachowicz & Foltran (2006) for Brazilian Portuguese.

3.1.1 Results

The results are summarized in Table 1 and Figure 2.

Table 1: Weak and generic definites and aktionsarten corpus occurrence (%)

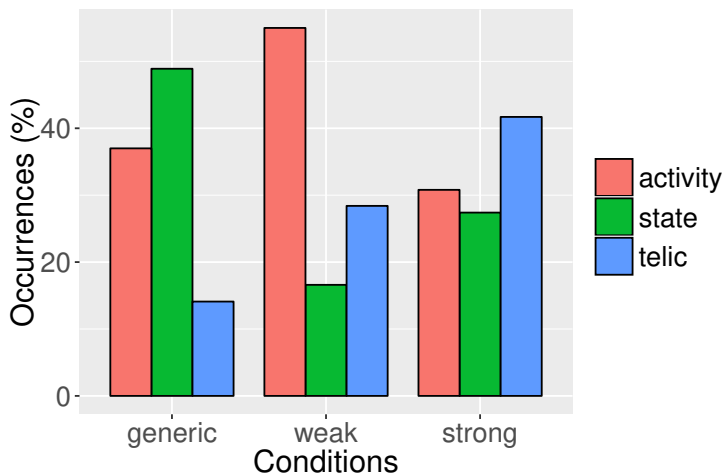| Conditions	Aktionsarten	Corpus occurrence (%)
Generic	State	48.9
	Activity	37
	Telic	14.1
Weak	State	16.6
	Activity	55
	Telic	28.4

Figure 2: Aktionsarten occurrence percentage in Generic, Weak and Strong conditions

Weak definites showed a significant difference (χ^2 = 171.6676, df = 2, p < 0.001) among state, 16.6%, activity, 55%, and telics, 28.4%, with activity being the most frequent category. Generic definites also significantly differ (χ^2 = 85.2335, df = 2, p < 0.001) in occurrences of state, 48.9%, telic, 14.1%, and activity, 37%.

The aktionsarten analysis is consistent with the incorporation hypothesis, in that weak definites are more frequent as activity and telic verbs. Also, as expected, generics are more frequent as state verbs. One interesting finding is that weak and generic definites are not in a complementary distribution.

3.2 Corpus summary

The quantitative data presented in this corpus analysis introduces some interesting evidence about weak definites. Weak definites are more frequent than generic definites. Weak definites occur in subject position and they do so less frequently than in object or adjunct position. Another interesting fact about syntactic position is that there is no complementary distribution between weak and generic definites, which would have provided support for the generic hypothesis.

The analysis of lexical aspect again found no complementary distribution between weak and generic definites. Also, as expected by the incorporation hypothesis, the majority of weak definites occur in activity and telic clauses.

4 Experiments

We conducted four experiments in which we compared participant's production and comprehension for stimuli that were chosen to bias weak, regular and generic readings. Our goal was to examine whether weak definites and generics exhibited similar properties as would be predicted by the simple version of the generic hypothesis. All of the experiments used the same materials, described in §4.1. The experiments were conducted in American English, they were programmed in *JavaScript*, and used *Amazon Mechanical Turk*[7] by the software *Psiturk*.[8] We used the Mechanical Turk platform because it provides easy and fast access to participants, data collection is reliable, and results are similar to those obtained in laboratory-based experiments (cf. Mason & Suri 2012; Paolacci et al. 2010).

4.1 Materials

The experimental materials were 54 sentences divided in three groups containing a noun phrase with a definite article which had: a clear generic reading (Example 7), a clear regular reading (Example 8) and a weak reading (Example 9):

(7) *Henry Ford created **the bus** in his early years.*

(8) *James crashed **the bus** during the night.*

(9) *Linda took **the bus** to go to college.*

For all sentences, the target noun was presented in a definite noun phrase which was an object of a telic verb or an activity verb. In our examples, *bus* is the target word, it is preceded by *the*, a definite determiner *the bus*, in object position of a telic verb, as *created*, *crashed*, *took*.

In Example (7) the sentence in the target DP has a prototypical generic reading, in which *the bus* has a kind uniqueness (cf. Carlson & Pelletier 1995; Carlson 2006). In Example (8), the *the bus* has a unique referent in the sense of Russell (1905). In Example (9), the DP supports a weak definite reading. The weak definite sentences were modeled on examples from Carlson & Sussman (2005); Carlson et al. (2006; 2013) and Klein et al. (2013).

---

[7]Access on: https://www.mturk.com/mturk/welcome
[8]Access on: https://psiturk.org/

The 54 sentences were divided into 3 lists of 18 sentences, each list with six exemplars of each type: regular, generic and weak. The same noun was never repeated within a list. The same noun appeared in a different condition in each list. Each participant was presented with one of the lists.

We briefly describe each of the four experiments in the following subsections.

4.2 Experiment 1: Judgment

The first experiment used a judgment task in which the participants judged whether the DP referred to either an individual or a category. We reasoned that regular definite noun phrases would be rated as referring to individuals whereas generics would be rated as referring to categories. Finding this pattern would provide important evidence that we had successfully created a set of materials with regular reference and a set with generic reference. The critical question was whether weak definites would pattern with the generics, as suggested by the generic hypothesis, or with regular definites. Participants read one sentence on each trial and judged if the bold word (the target word in one of the readings) was either a *CATEGORY* or a *INDIVIDUAL*, using a continuous scale, ranging from 0 to 100 with the words *INDIVIDUAL* and *CATEGORY* as the endpoints. Whether the first endpoint was individual or category was balanced within lists, as showed in the Figure 3.

```
Linda took the bus to go to college.
```
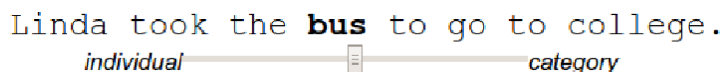*individual* ———————— ▣ ———— *category*

Figure 3: Judgment task screen – Sentence with the word *bus* to be evaluated on a continuous scale (screenshot)

We expected that the noun with a regular reading would be judged as an *individual* while the generic would be evaluated as *category*. This pattern of results is necessary to validate the task. The generic hypothesis predicts that the weak definites should pattern with the generic definites, as we can see in Table 2.

Table 2: Judgment task – Hypothesis according to generic theory

Definite readings	Weak = Generic
Generic	*Category* judgment (uniqueness of a kind)
Regular	*Individual* judgment (uniqueness)
Weak	*Category* judgment (uniqueness of a kind)

4.2.1 Participants

90 workers (40 women) from MTurk (https://www.mturk.com/) participated for payment of US$0.30. All participants provided informed consent in this experiment and in all of the other experiments we report.

4.2.2 Results

We analyzed the data using a Linear mixed model fit by REML ['lmerMod']. Using 0 as the individual endpoint and 100 as the category endpoint, regular definites were rated as closest to individual endpoint (mean = 19.82), whereas generics were rated as closest to the category endpoint (mean = 80.63). Weak definites were rated as closer to the individual endpoint (mean = 34.56). However, they fell between the regular and generics (Figure 4). Importantly, weak definites differed significantly from both the regular and generic noun phrases (Table 3).
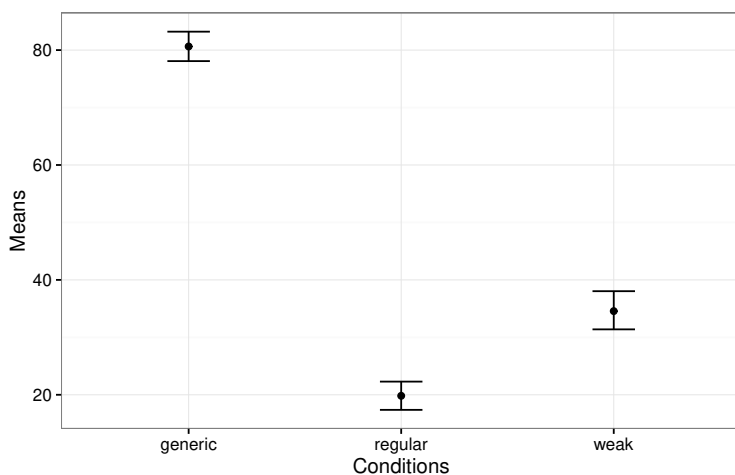

Figure 4: Judgment task – Judgment means (individual to category) by condition

Table 3: Judgment task – Statistics – Linear mixed model fit by REML ['lmerMod']

Formula:			
`ScaledResponse ~ condition + (1 + condition | subject) + (1 |, item),`			
Data:			
`data, Control: lmerControl(optimizer = "bobyqa")`			
	Estimate	Std. Error	t-value
(Intercept)	80.561	2.756	29.24
Regular condition	−60.719	4.501	−13.49
Weak condition	−46.133	4.134	−11.16

The results provide clear evidence that we successfully created two sets of sentences using the same nouns, that when used with a definite article in a DP, had a regular reading for one set and a generic reading for the second set. This serves as important validation for the materials. We also tested the prediction that if weak definites are, in fact, generics then they would show the same pattern. However, the sentences with weak definite noun phrases did not pattern with generic noun phrases and they were more similar to regular definite noun phrases than they were to generics. We note, however, because judgments of weak definites fell between the regular and the generics, one could argue that weak and generic definites are not different. One characteristic of noun phrases that have weak definite readings is that they can also be interpreted as regular definites. Therefore the results for the weak definites could, in principle, reflect a mix of regular and generic interpretations.

One way to assess the mixture possibility is to examine the distribution of responses to the three types of stimuli. If weak definites were a mix of regular and generics, we might expect to see a bimodal distribution, with an increased number of responses near the category endpoint. Figure 5 shows the distributions. Inspection of the patterns does not seem to support for the mixture hypothesis. Nonetheless this remains a possibility for results in which weak definites are intermediate between regulars and generics.
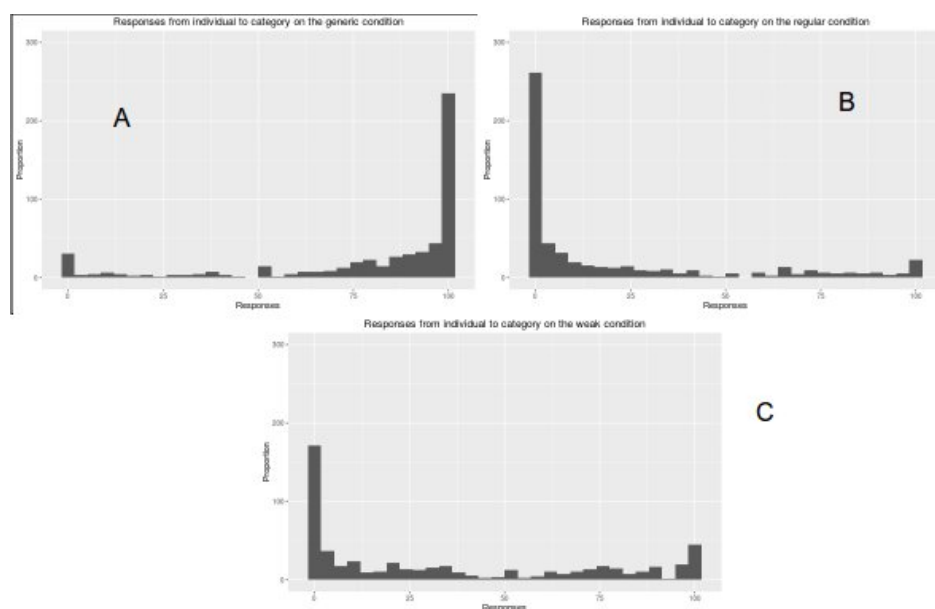
Figure 5: Judgment task – Condition histograms: (A) Generic distribution, (B) Regular distribution, (C) Weak distribution

4.3 Experiment 2: Forced choice

Our second experiment used a forced choice task, in which participants were presented with the same sentences as those use in the previous experiment. Participants were asked to choose between two possible noun phrases for a continuation sentence. One was a noun phrase that was anaphoric with the definite noun phrase in the preceding sentence (e.g. *That telephone...*). The other was a noun phrase that would introduce a new referent (e.g. *A telephone...*) (Figure 6).

Frank answers the telephone promptly.

A telephone...   That telephone...

Figure 6: Forced choice task screen

Our rationale was that regular definites would most likely be interpreted as referring to an individual, therefore licensing an anaphoric reference. In contrast the kind-reference supported by a generic would be more consistent with a continuation that introduced a novel referent. If weak definites are indeed a kind of generic, we would expect subjects to choose a new referent more often than the anaphoric continuation, i.e. weak definites would behave more like generic ones.

4.3.1 Participants

We again tested 90 workers (34 women) from MTurk for a payment of US$0.30, using the same lists as those created for Experiment 1.

4.3.2 Results

Figure 7 and Table 4 show the results. As we can observe, in sentences with the generic definite participants preferred a new referent (76.7%), while the regular reading showed the opposite preference, with 23.4% new referents. The weak definite did not pattern with the generic, participants chose a new referent only 42.9% (Table 5).

Results confirmed the expected pattern both for the clearly generic and regular expressions. Although the weak definites did not pattern with the generics, they showed fewer anaphoric choices than regular definites. This is not surprising because on the one hand, weak definites do not require a uniquely identifiable referent but, on the other hand, a weak definite noun phrase can easily be shifted to an interpretation with a uniquely identifiable referent.

Again however, one could argue that the results for weak definites could reflect a mix of generic and regular definites, In order to provide more nuanced evidence that did not require a meta-linguistic judgment with a binary choice, we conducted two production experiments.

4.4 Experiment 3: Free completion

In this experiment participants generated continuations for the sentences used in the previous experiments. No specific constraints were put on the form of the continuations except that participants should not use language that would upset their grandparents, as in Figure 8.

```
The great German composer, Wagner, changed the opera
                     for good.
```

Figure 8: Free completion task screen

We analyzed the continuations to see if they repeated the definite expression. The logic of the analysis was based on the incorporation hypothesis by Carlson et al. (2013) and Klein et al. (2013). If weak definites are indeed part of incorporated structures, then the event would be more salient than an individual referent would be introduced by a regular definite noun phrase or a kind-reference as introduced by a generic.
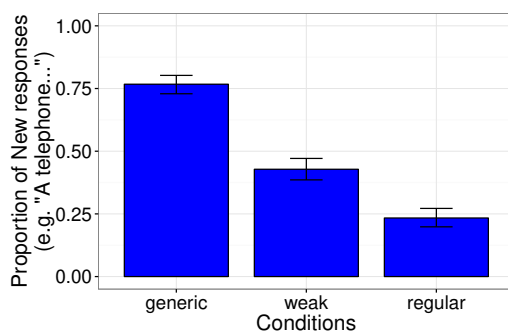
Figure 7: Forced choice task – Proportion of NEW by condition

Table 4: Forced choice task – Proportion of NEW and OLD by condition

Conditions	New (*A X…*)	Old (*That X…*)
Generic	0.767	0.233
Regular	0.234	0.767
Weak	0.429	0.571

Table 5: Forced choice task – Generalized linear mixed model fit by maximum likelihood (Laplace approximation) ['glmerMod']

Family:
```
binomial, ( logit )
```
Formula:
```
choice == "New" ~ condition + (1 + condition | subject) + (1 |,item),
```
Control:
```
glmerControl(optimizer = "bobyqa")
```

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	1.4719	0.2118	6.949	3.67×10^{-12}	***
Regular condition	−3.3248	0.3485	−9.540	$<2 \times 10^{-16}$	***
Weak condition	−1.9284	0.3130	−6.162	7.18×10^{-10}	***

4.4.1 Participants

90 workers (55 men) from MTurk participated for the payment of US$3.00.

4.4.2 Results

The frequency of repetition of the target word (e.g. *opera*) by condition was evaluated. The continuation in (10) is an example[9] of a situation which there was no target word repetition; the experimental sentence had the target word *opera* that was not used in the completion.

(10) Experimental sentence: *The great German composer, Wagner, changed the opera for good.*
Completion: *He was a beautiful person.*

We considered as repetition occurrences in which the target word was repeated in a pronoun form, as a DP (any kind of determiner + target word) or as a bare noun (only the target word, on in either plural or singular form). In Example (11), the repetition by a pronoun form (i.e. *it*) can be observed. In Example (12), the DP repetition occurred (i.e. *the opera*). The last example, (13), shows bare noun repetition (i.e. *operas*).

(11) Experimental sentence: *The great German composer Wagner changed the opera for good.*
Completion: *It is now much better then before.*

(12) Experimental sentence: *The great German composer Wagner changed the opera for good.*
Completion: *The opera is still a noble entertainment today.*

(13) Experimental Sentence: *The great German composer Wagner changed the opera for good.*
Completion: *Many later operas incorporated his changes.*

Table 6 and Figure 9 show that our hypothesis was confirmed, the weak definite was significantly less repeated (see Table 7 for stats) than the other definite conditions.

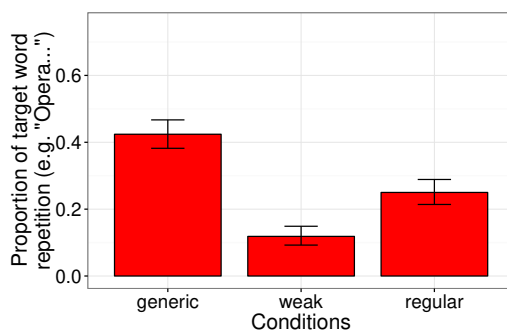[9]All the following examples are from data.

Figure 9: Free completion task – Proportion of target word repetition by condition

Table 6: Free completion task – Proportion of target word repetition (YES) and no-repetition (NO) by condition

Condition	NO	YES
Generic	0.576	0.424
Regular	0.750	0.250
Weak	0.881	0.119

Table 7: Free completion task – Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family:

```
binomial, ( logit )
```

Formula:

```
twr == "y" ~ condition + (1 | subject) + (1 |,item)
```

Data:

```
datac
```

Control:

```
glmerControl(optimizer = "bobyqa")
```

	Estimate	Std. Error	z-value	P(>|z|)	
(Intercept)	−2.3662	0.2654	−8.915	<2 $\times 10^{-16}$	***
Weak condition	1.9978	0.3454	5.785	7.27×10^{-9}	***
Regular condition	1.0323	0.3487	2.960	0.00308	**

The results showed that the definite noun was more likely to be repeated in a continuation for the generic and regular sentences compared to sentences with weak interpretations. Unlike the previous studies where the weak definites fall somewhere between regular definites, the regular and generics were similar to one another with the weak definites showing the fewest repetitions.

Moreover, when participants chose continuations with repetitions they tended to use different morphosyntactic forms and they made different semantic choices. As we can see in the occurrence examples below, (14–16), the experimental sentence has its target word in the generic condition in which *opera* is a kind. When the subjects repeated *opera*, they used three different morphosyntactic forms, but they kept the kind reading.

(14)   Experimental sentence: *The great German composer Wagner changed the opera for good.*
Completion: **It** *is now much better then before.*

(15)   Experimental sentence: *The great German composer Wagner changed the opera for good.*
Completion: **The opera** *is still a noble entertainment today.*

(16)   Experimental sentence: The great German composer Wagner changed the opera for good.
Completion: *Many later **operas** incorporated his changes.*

The morphosyntactic choices for generics was interesting, especially the use of bare noun forms, which have a generic reading. The final experiment used a forced completion task to investigate the forms that repetition would take.

4.5 Experiment 4: Forced completion

Another group of participants was asked to generate completions. In contrast to Experiment 3, participants were instructed to repeat the bolded noun used in the first sentence. However, they were not given any instructions about the form of the repetition.

The determiner choice (bare, definite, pronoun) was analyzed. We expected that, if in the first sentence there was a generic definite expression, then participants would be more likely to use the noun in a bare plural expression compared to a regular definite. Taken as a whole, the pattern of results from the previous

experiments would suggest that weak definites would show similar patterns as regular definites, with minimal use of bare nouns.

```
In the XVIII century, hygiene rules were introduced
       into the hospital in the western world.
```
Submit

Figure 10: Forced completion task screen

4.5.1 Participants

30 workers (16 men) from MTurk participated for the payment of US$3.00.

4.5.2 Results

In all conditions the definite article + noun ("dp" in Figure 11) was the most used form of repetition, as expected, both because the definite expression was used in the first sentence and because it is by most frequent kind of nominal phrase. However, bare plurals were sometimes used, but only in the generic condition ("bp" in Figure 11). In fact it was the the second most preferred repetition form for the continuations following generic sentences. Crucially bare plurals were never used in continuations that followed weak definites.
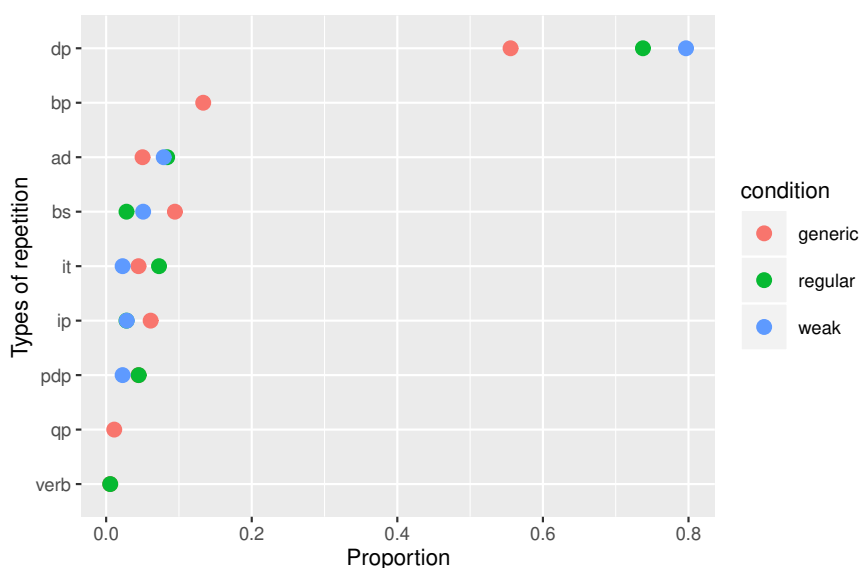

Figure 11: Forced completion – Types of repetition by condition

Below there are some completions and examples of some different morphological forms of repetition founded in our data. Example (17) is a "dp" (the+noun) occurrence; example (18) a "bp" (bare plural noun); example (19) an "ad" (noun transformed into an adjective); example (20) a "verb" (noun transformed into verb).

(17)   Experimental sentence: *In the XVIII century hygiene rules were introduced into the hospital in the Western world.*
Completion: ***The hospital*** *was now a clean place.*

(18)   Experimental sentence: *In the XVIII century hygiene rules were introduced into the hospital in the Western world.*
Completion: ***Hospitals*** *had never understood the importance of cleanliness.*

(19)   Experimental sentence: *In the XVIII century hygiene rules were introduced into the hospital in the Western world.*
Completion: *The* ***hospital*** *industry is now one of the largest in the world.*

(20)   Experimental sentence: *In Medieval times merchants used the bank to deposit their credit.*
Completion: *Merchants did a lot of* ***banking*** *and made money.*

Also in our data was the "bs" (bare noun singular), as Example (21); the pronoun (*it*), Example (22); the "ip" (noun determined by an indefinite article), Example (23); the "pdp" (noun determined by a pronoun), Example (24); the "qdp" (noun determined by a quantifier), Example (25).

(21)   Experimental sentence: *Most songwriters use the guitar when writing songs.*
Completion: ***Guitar*** *is the perfect instrument to work out music.*

(22)   Experimental sentence: *Samuel sold the guitar last year.*[10]
Completion: *He didn't want to sell* ***it*** *because it was his favorite guitar but he needed the money.*

---

[10] *Samuel vendeu a guitarra no ano passado.*

(23) Experimental sentence: *Jimi Hendrix played the guitar better than anyone else.*
Completion: *Nowadays **a guitar** that was played by him is worth very much.*

(24) Experimental sentence: *Zack listens to the radio while he drives.*
Completion: *His car **radio** is an aftermarket system.*

(25) Experimental sentence: *In the XVIII century hygiene rules were introduced into the hospital in the Western world.*
Completion: ***Every hospital** since then uses the same rules.*

The morphosyntactic repetition form was another interesting finding which distinguishes weak and generic definites. Bare plural nouns only happened in generic condition, behaving differently from weak definites once again.

4.6 Summary of experimental findings

In sum, we created a set of materials in which we would compare the properties of weak regular and generic sentences with object DP. Experiment 1 established that the regular and generic sentences showed the expected properties with regulars being judged as being about an individual and the generics as about a category. The weak definites behaved more similarly to the regular definites than the generics. In Experiment 2 we found that, as expected, regular definites licensed anaphoric completions, whereas generics encouraged interpretations that introduced new events. Again weak definites behaved more similarly to regulars compared to generics. Experiment 3 found similar results in a free completion task. Finally, Experiment 4 required participants to repeat the noun phrase in their completions, the distribution of the completions, suggested that generics behaved differently from both regular and weak definites.

5 Conclusions

In this chapter we presented new data from a corpus analysis and a set of experimental studies that examined properties of weak definites, regular definites and generics. The goal of this work was to provide additional evidence that could be used to evaluate the hypothesis that weak definite noun phrases are in fact generic DP.

In a corpus analysis we found that weak definites and generics are not in complementary distribution in either the syntactic environments in which they appear on the semantic types of events as indexed by the verb. Moreover, as predicted by the incorporation analysis, the majority of weak definites occurred in activity and telic clauses, while generic definites occurred more frequently in state and activity clauses. In a set of experiments we first created and validated properties of regular, generic and weak definites. We found that for the most part, weak definites behaved more like regular definites than generics. We also evaluated the possibility that the behavior of weak definites could be accounted for by the hypothesis that the behavior of weak definites reflected a mix of trials in which the weak definite was given a regular definite interpretation and trials in which it was given a generic interpretation. This type of model was, however, inconsistent with the results of several of the experiments. In sum, then, we found little evidence to support the hypotheses that weak definites showed similar properties to generics.

Our results are consistent with the incorporation hypothesis in that it assumes that the non-uniqueness of reference in weak definites does not arise because it is a form of generic. Therefore it would have been problematic for the incorporation hypothesis if weak definites had, in fact, patterned with generics in our studies. Further research will be necessary to determine whether the absence of generic-like behavior in these studies would be consistent with the type of analysis argued for in Aguilar-Guevara (2014), which accounts for non-uniqueness by assuming that weak definites derive their non-uniqueness of individual reference by virtue of their generic status and their eventive properties by virtue of the KLR rules, described in detail in Aguilar-Guevara (2014). Addressing these issues is beyond the scope of the current chapter.

Although the results we presented and the linguistic phenomena that we discussed lead us to conclude that the semantic incorporation hypothesis provides an account of the behavior of weak definites without assuming that they are generics, it is important to conclude with some caveats. First in the corpus analysis weak definites frequently appeared in subject position, which is unexpected in the incorporation analysis. Secondly, the conclusions from our experiments bring evidence to bear on the two analyses only insofar as we have been able to tap into the relevant referential behavior with our tasks. Third, there are properties of weak definites, in particular the parallel about restrictions on modifiers for weak definites and generics, that receive a straightforward account on the generic analysis developed by Aguilar-Guevara (2014), but require additional work to be explained by the incorporation analysis. Fourth, the arguments for the

role of the definite article depend on the scoping analysis we presented, which has some precedents in the literature but is not addressed in these empirical studies. If this analysis proves problematic, it will be important to explore other alternatives. Finally, we want to emphasize a point that has emerged from the work that the authors have conducted in collaboration with each other and with other colleagues. For a phenomenon such as weak definites which involve subtle interactions between putative structures and conceptual representations, and for which the linguistic data is less than definitive, experimental studies that target particular hypotheses can prove to be an important complement to linguistic argumentation.

Acknowledgements

Abbreviations

BP	Brazilian Portuguese	bs	bare singular noun
bp	bare plural noun	it	pronoun *it*
DP	Definite Phrase	ip	indefinite article (*a*/*an*) + noun
dp	definite article (*the*) + noun	pdp	noun determined by a pronoun
ad	noun transformed into an adjective	qp	noun determined by a quantifier
		verb	noun transformed into a verb

References

Aguilar-Guevara, Ana. 2014. *Weak definites: Semantics, lexicon and pragmatics* (LOT Dissertation Series 360). Utrecht: LOT.

Aguilar-Guevara, Ana & Joost Zwarts. 2011. Weak definites and reference to kinds. *Proceedings of SALT* 20. 179–196.

Aguilar-Guevara, Ana & Joost Zwarts. 2013. Weak definites refer to kinds. *Recherches linguistiques de Vincennes* 42. 33–60.

Carlson, Greg N. 2006. Generic reference. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 14–18. Amsterdam: Elsevier.

Carlson, Greg N., Natalie M. Klein, Whitney Gegg-Harrison & Michael K. Tanenhaus. 2013. Weak definites as a form of definiteness: Experimental investigations. *Recherches linguistiques de Vincennes* 42. 11–32.

Carlson, Greg N. & Francis Jeffry Pelletier (eds.). 1995. *The generic book*. Chicago: University of Chicago Press.

Carlson, Greg N. & Rachel Shirley Sussman. 2005. Seemingly indefinite definites. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives* (Studies in Generative Grammar 85), 71–86. Berlin: Mouton de Gruyter.

Carlson, Greg N., Rachel Shirley Sussman, Natalie M. Klein & Michael K. Tanenhaus. 2006. Weak definite noun phrases. *Proceedings of NELS* 36(1). 179–196.

Clark, Herbert H. & Catherine K. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of discourse understanding*, 10–63. Cambridge: Cambridge University Press.

de Sá, Thaís Maíra Machado, Maria Emília Saramago & Maria Luiza Cunha Lima. 2016. A corpus data of weak definites in Brazilian Portuguese. *Revista da ABRALIN* 15(1). 101–120.

Donnellan, Keith. 1966. Reference and definite descriptions. *The Philosophical Review* 75(3). 281–304.

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

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts. (Doctoral dissertation).

Klein, Natalie M., Whitney Gegg-Harrison, Greg N. Carlson & Michael K. Tanenhaus. 2013. Experimental investigations of weak definite and weak indefinite noun phrases. *Cognition* 128(2). 187–213.

Mason, Winter & Siddharth Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behav Res Methods* 44(1). 1–23.

Paolacci, Gabriele, Jesse Chandler & Panagiotis G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5(5). 411–419.

Poesio, Massimo. 1994. Weak definites. *Proceedings of SALT* 4. 282–299.

Russell, Bertrand. 1905. On denoting. *Mind* 14(56). 479–493.

Strawson, Peter F. 1950. On referring. *Mind* 59(235). 320–344.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160.

Wachowicz, Teresa Cristina & Maria José Foltran. 2006. Sobre a noção de aspecto. *Cadernos de Estudos Llingüísticos* 48(2). 211–232.