

# Bad Subject: (Non-)canonicity and NP Distribution in Existentials

David Beaver, Itamar Francez & Dmitry Levinson  
*Stanford University*

## 1. Introduction

Constructions of the type exemplified by (1) have been extensively studied in linguistics and philosophy. In the generative literature they are generally referred to as “existentials”.

(1) There is a problem with this mobile phone.

One of the most extensively studied aspects of these constructions is the distribution of NP/DP<sup>1</sup> types that occur in the role or position of *a problem* in (1). Following common practice, we refer to the NP occurring in this role/position as *pivot*. Despite extensive study, there is no consensus as to what kinds of NPs occur as pivots, or why. In this paper, we argue:

- a. That the distribution of NP types in pivot role is more complex than has been acknowledged in the literature.
- b. That definiteness effects should be accounted for not by construction-specific constraints, but by general markedness constraints, in particular markedness constraints on grammatical subjects.
- c. That definiteness effects of the kind exhibited by existentials vary systematically across languages, and that understanding the nature of this variation allows us to identify a natural class of constructions cross-linguistically.

In section 2 we discuss the relation between existential and non-existential constructions and outline an informal theory of pivots based on markedness constraints on subjects. We also present qualitative data from English, Russian and Hebrew that demonstrates systematic variation in pivot properties across languages. In section 3, we present quantitative data from a cross-linguistic corpus study, which demonstrates that the distribution of pivots in various languages corresponds to the expectations arising from our theory. Section 4 contains discussion of two phenomena in Russian, the genitive of negation and the optionality of the copula *jest'* in present tense locatives and possessives. We show that neither of them involves a distribution like that standardly expected for definiteness effects, but that on our classification the first involves a clear definiteness effect, while the second shows a weaker effect.

## 2. The Marked Status of Existential Constructions

In many languages, the construction corresponding to the one exemplified by the English sentence in (1) deviates from the canonical structure of a declarative sentence in the language. Typological studies (e.g. Freeze 1992, Clark 1978) have pointed out various ways in which this deviation can be realized, including:

**Word order** Existential sentences often show a reversal of the canonical ordering of a theme NP and a locative phrase. For example, in Russian a canonical locative predication consists of a theme NP followed by a locative phrase (2). This corresponds to the general canonical form of predication in Russian, as seen in (3). In the existential on the other hand the canonical order is reversed, and the locative phrase precedes the subject, as seen in (4). A similar situation holds in Finnish and many other languages.

- (2) *kniga na stole* [RUSSIAN]  
 book on table  
 'The book is on the table.'
- (3) *Alfred kompozitor* [RUSSIAN]  
 Alfred composer  
 'Alfred is a composer.'
- (4) *Na stole jest kniga* [RUSSIAN]  
 On table COP book  
 'There is a book on the table.'

**Special copulas** In many languages, the existential construction involves a copula or verb that is distinct from the copula used in canonical constructions. For example, in the Russian examples above, the copular element *jest* 'is' occurs in existential constructions but not in general in copular constructions. Similar special copulas are found e.g. in Hebrew and Turkish.

**Locative pro-forms** In many languages, a locative deictic lexeme is grammaticalized into the existential construction, where it is stripped of its deictic function. Examples are French *y* 'there', and English *there*.

**Morphosyntax** Across languages it is common for existential constructions to show non-canonical and unstable morphosyntactic patterns of argument realization and agreement. Pivot NPs often show unstable case marking in that they fluctuate between canonical subject marking and marking more typical of objects.<sup>2</sup> For example, in both modern Hebrew and English, pivots fluctuate between nominative and accusative case. Also, existential constructions often exhibit fluctuation between agreement and no agreement (i.e. impersonal agreement on the existential verb/copula).

The relation between existential constructions and copular constructions has featured extensively in syntactic research, where the two constructions are often seen as derivationally related, or as having a common underlying structure. Research on the distribution of pivots, however, has not in general exploited this relationship. Copular and existential constructions are often similar in (truth-conditional) meaning, differing only in the way the NP argument is realized. Thus, while canonical subjects precede the copula, pivot NPs tend to follow the copula or existential verb. While canonical subjects precede locative phrases, pivot NPs tend to follow them.

A natural interpretation of this situation is that speakers choose to realize NPs as pivots when the unmarked option of realizing them as subjects in copular constructions is for some reason unavailable. The reason for its unavailability may relate to properties of the pivot NP itself, its compatibility with a canonical construction. We hypothesize that canonical copular constructions and existential constructions are in competition, and that this competition is directly related to the distribution of pivots.

- (5) **Markedness hypothesis** The opposition between existential and canonical structures is what underlies the distribution of NPs in pivot function cross-linguistically.

Below we suggest that pivot NPs are NPs that are not good candidates for functioning as subjects, for lack of prototypical subject properties.

### 2.1. *The “Definiteness Effect”*

It is well known that not all NP types are equally felicitous as pivots. Among the kinds of NPs that have been said not to occur as pivots are definite NPs as in (6) and (7), and strongly quantified NPs as in (8). These restrictions are known as the “definiteness effect”.

- (6) ??There is the problem (again) with this mobile phone.  
 (7) ??There is him on the couch.  
 (8) ??There is every cigar in David’s drawer.

The body of research dealing with the definiteness effect (see e.g. Keenan (2003), Barwise and Cooper (1981), Milsark (1977), McNally (1998), Zucchi (1995)) has mostly been based on constructed English examples. Consequently, the cross-linguistic variation in the extent of these restrictions on pivots has not been studied systematically, nor has it been taken into account in formulating theories of the definiteness effect.

While some authors have noted that the definiteness effect does not hold equally for all languages (see e.g. Ziv (1982) for Hebrew, Moro (1994) for Italian and Bentley (2002) for Sardinian), these authors do not present a theory of the definiteness effect that explains cross-linguistic variation. Similarly, even though it has

been noted that even in English, most NP types that have been claimed not to occur in existential constructions do in fact occur (with various interesting consequences for their interpretation), available semantic theories of the definiteness effect treat it as a categorical ban on a class of NPs, and focus mostly on trying to characterize this class. With the notable exception of Ward and Birner (1995), there have not been extensive attempts to explain the occurrence of definites and other NP types in English existential constructions in terms of general linguistic principles.

Mikkelsen (2002) develops a theory of Danish expletive constructions that is rooted in exactly this intuition. Mikkelsen's analysis treats the definiteness effect as epiphenomenal to the interaction of markedness constraints on subjects. She shows that no special constraints need to be posited for expletive constructions. Instead, harmonic alignment constraints built on well observed typological tendencies suffice to derive the distribution of pivots in Danish as well as English. The prototypical subject property that figures in her analysis is definiteness. Across languages, definite subjects are preferred over indefinite ones (Givon 1978). In Mikkelsen's theory, the variation in relative ranking of constraints against expletive subjects and against indefinite subjects derives the distribution of NPs in Danish and English.

Definite NPs however are only a subset of the NP types that have been said not to occur in existential constructions. This is not surprising on the present view, since other than definiteness, there are many other properties that characterize canonical subjects, and their importance varies from language to language. While some of them are grammaticalized in some languages, others produce statistical tendencies. We aim to show that the competition driving the distribution of definite NPs and expletives in Mikkelsen's theory should be extended to all NP types.

The idea that markedness can be captured by associating different hierarchical scales goes back to work by Michael Silverstein, and is by now well established in the typological/functional literature, especially in Optimality Theory. Silverstein (1976) proposed that typologically, elements that are high on a person/animacy hierarchy are by default associated with more prominent thematic roles. Aissen (1999) provides the following version of Silverstein's hierarchy:

- (9) Person/Animacy hierarchy:  
 a. Local person > Proper Noun 3rd > Human 3rd > Animate 3rd  
 > Inanimate 3rd  
 b. Agent > Patient

Aissen implements Silverstein's ideas formally in OT, and shows how his general typological idea applies in particular grammars. In existential constructions, we take it that the kind of scale offered by Silverstein and Aissen serves in parallel to other scales such as the definiteness scale discussed by Mikkelsen, as well as other scales yet to be precisely defined. For example, in the literature on existential constructions it is commonly observed that use of the construction is influenced by discourse pragmatics, and that pivots tend to differ pragmatically from subjects. Ward and Birner (1995) claim that existential sentences function to introduce new discourse referents, and hence that pivots are hearer new. Relatedly, Reinhart (2005) among others has claimed that pivots cannot be sentence topics,

while Zucchi (1995) claims that they cannot be presuppositional. It is well known that canonical subjects are often discourse topics and hence both hearer old and presuppositional. On a scale based approach, these observations would be captured by employing a pragmatic scale such as the one in (10).

- (10) Topicality scale:  
discourse topic > discourse old > hearer old > hearer new

Such a scale is in line with properties that have been independently claimed to correlate with subjecthood in the literature, e.g. by Keenan (1976), and are known to be involved in phenomena sensitive to subjecthood such as voice alternations.

In this paper we do not provide a scale of subject properties *per se*. Rather, we argue for orderings of NP types that are plausibly related to underlying scales of subject properties. While the definiteness effect is standardly seen as categorical, on the scale based view presented by Mikkelsen and adopted and expanded here, it is inherently more complex. The definiteness effect is the product of a kind of optimization relative to subject properties:

- (11) **Subject properties hypothesis:**  
NPs that exhibit properties associated with subjecthood are attracted to constructions involving a canonical subject, whereas those that do not display such properties are attracted to constructions that do not involve canonical subjects.

If the picture just outlined is correct, then we expect three things:

**Gradient distribution:** Since there are usually many factors affecting the likelihood that a certain NP type will be subject, we expect different NPs that would traditionally fall into one type, e.g. strong NPs or weak NPs (Milsark 1977), to vary in their propensity for occurring in either of the two relevant constructions. In other words, the definiteness effect should be gradient rather than categorical.

**Anti-definiteness effects:** If NPs that do not occur often in existential constructions (the definiteness effect) are prototypical subjects, then we expect NPs that do occur often in existential constructions, i.e. the pivots, not to be prototypical subjects. The definiteness effect has a flip side, in that certain NP types should be restricted in their capacity to occur as canonical subjects. These NPs should then be “attracted” to the pivot position. This follows from the logic of markedness: NPs that *can* be realized as canonical subjects will be so realized.

**Systematic variation:** We not only expect languages to vary in the degree to which they show a definiteness effect (due to differences in which properties count towards canonicity of subjects), but we also expect that variation to be in a consistent direction. For example, we expect that two NP types that are ordered in relation to their propensity for subjecthood will not switch their

order in another language. Similarly, we expect that there are no languages in which strongly quantified NPs are much worse subjects and much better pivots than weak indefinites. Our theory therefore suggests implicational universals.

From an empirical point of view, such an account enables a refined diagnostic for deciding when a construction exhibits a definiteness effect: a construction involves such an effect if the distribution of NPs in it is correlated with the ordering of NP types, seen as an implicational universal. In section 4 we discuss the use of such criteria with reference to Russian NP distribution.

As mentioned earlier, the gradience of the distribution of NP types in existential sentences and the systematicity of this gradience have not been noted in the literature. The observation that indefinites are infelicitous as subjects under certain interpretations (but not others) is on the other hand rather old. In particular, as Mikkelsen (2002:25) notes, it has long been observed that so called ‘weak’ indefinites are bad subjects cross-linguistically, and in fact are ungrammatical as subjects in some languages on a non-generic interpretation. However, this fact about the distribution of indefinites has not to our knowledge been related in any way to the definiteness effect in existentials.

## 2.2. Evidence for Cross-linguistic Variation in Existentials

We now present English, Russian and Hebrew data that provides initial evidence of cross-linguistic variation, and we show that the type of variation we observe is consistent with the hypotheses suggested above. The data we consider involves pronominal elements, and centers on three factors: the locality of pronouns, the referentiality of pronouns and the polarity of the construction. Based on these factors, Russian, Hebrew and English form a cline with respect to pronoun types allowed in existential constructions.

Consider first sentences (12) - (14) as answers to a question like *where is my wallet?* All three sentences are very marked as answers to this kind of question:

- (12) ?? There’s it on the table.
- (13) ?? ona jest’ na stole. [RUSSIAN]  
       it COP on table  
       ‘There is it on the table.’
- (14) ?? yeš oto al ha-šulxan [HEBREW]  
       COP him on the-table  
       ‘There’s it on the table.’

Now consider the sentences (15) - (17), evaluated as answers to a question like *where can I find this book?* The English sentence is still very marked, whereas the Russian and Hebrew sentences are natural.

- (15) \* There's it in the library.
- (16) ona jest' v biblioteke. [RUSSIAN]  
it.NOM is in library.  
'They have it in the library.' (lit.: There's it in the library.)
- (17) yeš oto ba-sifriya. [HEBREW]  
COP him in-the-library.  
'They have it in the library.' (lit.: There's it in the library.)

The semantic interpretation of the pivot in (16)-(17) is different that that in (12)-(14). While the latter sentences are used to answer a question about a particular wallet, the former are used to answer a question about any token of the book under discussion. Loosely, we may say that the pivot in (16)-(17) is used less referentially. The exact nature of this difference in referentiality is not directly relevant to our current purposes, but we note that it is related to the distinction made by e.g. McNally (1998) between type and token denoting NPs. We thus refer to the NPs in (16)-(17) as type-denoting, and those in (12)-(14) as token-denoting. The relevant point here is that Hebrew and Russian are both much more permissive than English in allowing type-denoting pronouns in existential constructions.

The situation is different when polarity is changed. While inanimate pronouns are still very marked in English in negation contexts, in Russian and Hebrew negated contexts seem to remove the limitation of pivots to type-interpretation. Thus, (19) and (20) are good as answers to the question *where is my wallet?* :

- (18) \*There isn't it on the table.
- (19) jejo net na stole / na stole jejo net. [RUSSIAN]  
it.GEN not-is on table / on table it.GEN not-is.  
'It's not on the table.' (lit.: There isn't it on the table.)
- (20) eyn oto al ha-Sulxan. [HEBREW]  
COP him on the-table.  
'It's not on the table.' (lit.: There isn't it on the table.)

Now consider the sentences in (21)-(22), involving a local pronoun. These sentences are in all three languages just as marked as (12) - (14) when used as an answer to a question like *where are you?*:

- (21) ?? There is me at home.
- (22) ?? ja jest' doma. [RUSSIAN]  
I is home  
'There's me at home.'
- (23) ?? yeš oti ba-bayit. [HEBREW]  
COP me in.DEF-home  
'There's me at home.'

However, when the polarity is changed, there is a difference between Russian and Hebrew: while such sentences are not acceptable in Hebrew, they are acceptable and in fact common in Russian:

(24) \* There isn't me at home.

(25) Menja net doma. [RUSSIAN]  
I.GEN not-is home.  
'I'm not home.' (lit.: There isn't me home.)

(26) ??eyn oti ba-bayit. [HEBREW]  
COP.NEG me in.DEF-house  
'I'm not home.' (lit.: There isn't me at home.)

The data are summarized in table 1, which shows a systematic increase in the acceptability of pronouns: English is least permissive, Hebrew is less permissive than Russian, and Russian is most permissive. Importantly, note that properties associated with subjecthood nicely predict how acceptable a pronoun type is likely to be cross-linguistically. The more subject-like the NP, the less likely it is to be acceptable in existentials.

<i>Type of pronoun</i>		<i>Examples</i>	ENGLISH		HEBREW		RUSSIAN	
			<i>pos</i>	<i>neg</i>	<i>pos</i>	<i>neg</i>	<i>pos</i>	<i>neg</i>
Local		(21) - (26)	-	-	-	-	-	+
Non-local	token	(15) - (20)	-	-	-	+	-	+
	type	(12) - (14)	-	-	+	+	+	+

Table 1: Acceptability of pronouns in existentials in English, Russian, and Hebrew.

Arranging the pronouns according to how common they are in existentials cross-linguistically, we get that type denoting non-local pronouns are most common, then token denoting non-local pronouns, and then local pronouns, which only occur in Russian, and only under negation. This ordering is exactly what we expect if pivots are bad subjects. It corresponds (in reverse) to the following two subject-property scales:

Local Pronoun > non-local pronoun

NPs with token interpretations > NPs with type interpretations

That locality is a factor in subject selection is well known from typological literature,<sup>3</sup> and follows from local pronouns being highest on the animacy/prominence scale. Furthermore, this order emerges in the quantitative studies reported below, and the convergence of results is obviously not incidental. That token interpretations are more subject-like than type interpretations follows naturally from the assumption that subjects tend to be presuppositional and specific.

The table above supports our claim that the ordering of NP types in existentials constitutes an implicational universal: if a language allows local pronouns, it also allows non-local pronouns, and if it allows token denoting pronouns it also allows type denoting ones. Although a detailed theory is lacking, it is unsurprising that negation should affect pivot distribution, since negation alters the set of properties the subject is entailed to have. We predict that such asymmetry, when it exists, should consistently hold in the same direction. Independent polarity licensing facts apart, no NP types should be more natural as pivots in positive contexts than in negative ones. In the quantitative studies reported in the next section, a similar effect will be seen: downward monotone quantifiers, including the negative quantifiers *no* in English and *geen* ‘no’ in Dutch, will be shown to be more strongly existential than upward monotone quantifiers in a sense to be made precise.

### 3. Quantitative Data from Dutch and English

Most data on definiteness effects in the literature is drawn from judgments on constructed examples. We do not suggest dispensing with judgments on artificial data, but think it useful to supplement such evidence with quantitative data based on naturally occurring examples. Much of the data reported here is drawn from Google searches. However, many methodological issues come up in the use of data obtained from the web using standard search engines, so it was important to check that the results could be duplicated using search tools dedicated to linguistic work on structured corpora. We report on some preliminary results from these latter investigations below.

In the first two sections of this paper we suggested that the distribution of NPs as existential pivots be understood in terms of a competition: an NP can appear in pivot position when it lacks canonical subject properties. This suggests a new way of looking at definiteness effects in existential constructions. Standardly, the distribution is considered *in vacuo*: a given NP type either does or does not occur in existential pivot position. What we will consider instead is the prevalence of NP types in existential pivot position relative to their prevalence in canonical subject position.

#### 3.1. Method

The quantitative study involves the following variables:

**NP type** The classes of NP we consider are mostly either standard lexical NP categories (e.g. pronouns) or else classified according to the head determiner.

**Position** For the web-based studies reported here we considered two templates. The existential template for English is “there s|is|are [NP pattern] in”, and the canonical form template is “[NP pattern] is|are in”. Here [NP pattern]

might, for example, be “every \*” where \* is a wildcard matching any single word. Note that for the Google searches we generally restricted ourselves to NPs with a single noun, though in our structured corpus searches we dispensed with this constraint.

The web based searches used patterns including a preposition “in”. The reason is that simple searches based on a template “[NP pattern] is” produce a high percentage of matches which are not of the desired form [S [NP [VP [V be] ...]]]. Searches which are more specific, due to the inclusion of an additional element, produce a higher percentage of matches of the desired syntactic form. Our searches on structured corpora (all pre-parsed) did not require an explicit preposition since we could extract appropriate syntactic matches with a high success rate.

**Language** In this section we report on results for English and Dutch, and in section 4 will report on quantitative work on Russian. Others in our research group have performed quantitative analyses for Danish and Spanish, and we hope to report on these at a later date.

**Genre** For English we performed studies using corpora constructed from quite different sources. The web is heterogeneous, and has a low percentage of transcribed speech. The three structured corpora we used were: Switchboard, a speech corpus drawn from telephone conversations; Brown, a balanced corpus including a mixture of transcribed speech and other text types; and WSJ, a collection of newspaper articles from the Wall Street Journal.<sup>4</sup>

**Frequency** The only dependent variable is frequency, i.e. the number of matches. However, for web derived data the raw number of matches using standard web search tools is not adequate because of the high incidence of inappropriate matches. To offset this problem, for each separate search we sampled the results. Then, rather than using the raw number of search matches, we used adjusted frequencies, that is, the raw number of search matches multiplied by the proportion of search results which have the intended form as measured by our sample.

There are a number of ways to study the relative frequency of NP types in canonical and existential constructions. For example, it would be reasonable to consider for a given NP type the proportion of NPs of that type which occur in existential constructions out of the total number of tokens of that NP type in either construction. However, we opt for a different metric, namely the ratio of the number of canonical tokens to the number existential tokens. The downside of using this metric is that it is undefined in cases where we found zero occurrences of an NP type in existential pivot position, but the upside is that despite the small number of NPs for which the metric fails, it none the less brings out the differences between NP types more clearly than other statistics we considered. For cases where the adjusted number of existential pivot occurrences was zero, we (naturally enough) treated the ratio as tending to infinity.

The range of variation of the Canonical/Existential ratio is enormous, even for cases where it is well defined. In our English data, all NPs we considered had non-zero numbers of existential pivot occurrences, but the ratio ranges between 0.02 and 170,000. At the low end are NPs of the form *at most N* with a ratio of 0.0247. These are thus extremely existential, by which we mean that speakers have a strong preference for putting them in existential pivot position rather than canonical subject position when the choice is available. At the high end are local pronouns, with a ratio of 167,000, which are thus highly canonical.

To exemplify how the ratios are calculated, consider a simple example involving the NP type *few N*. We firstly ran searches on the quoted strings “*there s|is|are few \* in*” for existential occurrences and “*few \* are in*”. The \* matches exactly one word, and the |-operator is interpreted as disjunction.<sup>5</sup> For the existential construction, there were 105,000 hits, and for the canonical construction there were 34,000. We manually sampled the results of each search.<sup>6</sup> For each construction, we counted the percentage of cases where our analysis of each sample did not correspond to the structure we were seeking. We counted as errors any case where the syntax was not what we were seeking, i.e. *NP be-form PP* for the canonical construction and *expletive-there be-form NP PP* for the existential. In some cases, as when a sentence boundary intervened in the matched string, the classification was easy, but in other cases, as in deciding whether a complex expression is a PP or deciding whether an occurrence of *there* is expletive or locative, the judgement is subtle. For searches returning very small numbers of results ( $\leq 100$ ) we checked all results. For most searches the number of results was much higher (up to 10,000,000). For these we performed an initial sample of 20 results. If the error rate was above 50%, we then increased the sample size, sampling up to 100 results in the few cases where the searches returned close to 90% erroneous results (4 cases out of 46). In the case of *few*, 10% of the existential results were erroneous, and 75% of the canonical results were erroneous. One source of the high error rate for *few* in canonical constructions is that many of the results were in fact for *a few*.<sup>7</sup> Based on these error rates, we calculated adjusted frequencies. For *few* in existentials the adjusted figure is  $105,000 \times 0.90 = 94,500$ , and for *few* in canonical constructions it is  $34,000 \times 0.25 = 8,500$ . We then calculated the ratio of the adjusted canonical frequency to the adjusted existential frequency, which in this case yields a ratio of  $8,500/94,500 = 0.09$ .

As we see in the above example, existential occurrences of NPs of the form *few N* outnumber comparable canonical constructions by a factor of more than 10, and we can say that *few* is strongly existential. But such a claim also highlights an important limitation of the type of quantitative data we have obtained so far, a limitation that applies equally to all NP types we have looked at. The quantitative data we collected does not determine whether the 91% of existential occurrences of *few* have an intrinsically different meaning than the 9% of canonical occurrences. If it should happen that this split corresponds precisely to the standard distinction between proportional and cardinal uses of *few*, then that would imply that cardinal uses of *few* in locative copular constructions vastly outnumber proportional uses. However, standard semantic accounts of the definiteness effect suggest that while

the proportional use should always be strong, the cardinal use should allow for both weak and strong readings. It would follow that proportional uses of *few* are very rare indeed. Then again, perhaps those standard theories are incorrect. The methodological message we want to make explicit is that the data we present in this section simply does not determine whether proportional uses of *few* ever occur in existential pivot position, and similarly does not determine in general whether occurrences of a given NP form carry different meanings in canonical subject and existential pivot position.

### 3.2. Results

For the study of English NP distribution using Google, results are presented for a range of NP types in figure 1.

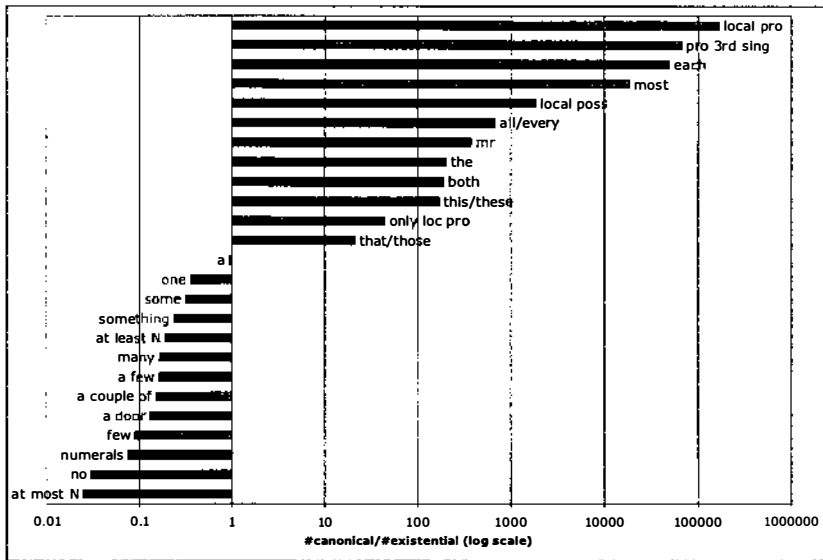


Figure 1: English Canonical/Existential ratios

#### *Result 1: No categorical definiteness effect*

The first thing to note about the results we obtained is that they provide no direct evidence for a categorical definiteness effect: all the NP types we searched for in English that are themselves reasonably frequent occur with some non-zero frequency in existential pivot position.

*Result 2: A strong non-categorical definiteness effect*

Result 1 does not imply that there is no definiteness effect. On the contrary, the results show a very strong definiteness effect, but it is an effect of a different sort than that usually observed. Looking at the table, we see that on the left hand side with canonical/existential (C/E) ratios between 0 and 1, are a range of mostly indefinite and downward monotone NP types. On the right hand side, with ratios between 20 and 200,000 are a mixture of definites and proportional determiners. That is, weak NPs appear on the left, and strong NPs on the right, with the notions of weak and strong falling on standard lines. But the definiteness effect, as it appears in our data, is not the categorical effect normally described. Rather, the definiteness effect consists of a massive gulf in the space of C/E ratios. Let us say that one NP is *more existential* than another if its C/E ratio is lower, and that one NP is  $\alpha$  times as existential as another if its C/E ratio is  $1/\alpha$  as big. Then for English, the least existential weak NP is 20 times as existential as the most existential strong NP.

*Result 3: Anti-definiteness effects*

The weak NPs vary in existentiality by a factor of over 20. Several determiners are strongly existential, notably *at most*, *no*, *few* and numerals, all of which favor existential pivot position by a factor of over 10 (a factor of 24 for *at most*): we describe such NPs as manifesting an *anti-definiteness* effect. Note that three of these heavily existential NPs are downward monotone. We speculate that their high existentiality relates to the fact that unlike upward monotone cardinal determiners they lack generic or specific readings and do not ordinarily introduce discourse referents. As regards numerals, we suggest that the high degree of existentiality we found may relate to their lacking generic readings and tending not to be given specific interpretations.<sup>8</sup>

Although simple indefinites headed by *a(n)* are often considered prototypical examples of determiners occurring in existential pivot position, in terms of C/E ratios they are not at all prototypical. In fact simple indefinites (with generic, specific and other readings lumped together) occur at approximately equal rates in canonical subject position and existential pivot position. With a C/E ratio approaching 1 in the Google data, simple indefinites cannot be said to exhibit any anti-definiteness effect (but see below for a discussion of genre variation for *a(n)*).

We defined most NP forms using the head determiner as a matching criterion, but the bar chart also includes the NP *a door*, which turns out to be strongly existential, favoring pivot position by a factor of 13:1. Two factors are relevant. First, *a door* often refers to an object which is part of some other element mentioned in the sentence, e.g. a room or a wall. Thus, *a door* refers to a dependent element, such that if you move the thing it depends on, the door moves too, while the reverse does not hold. Subjects, we suggest, are preferably independent, so that neither their position nor existence intrinsically relies on that of other objects referred to in the same clause. A second relevant factor is animacy: universally, subjects tend to be animate, so it may be that inanimates are disfavored in subject

position and tend to be more existential than similar animates. However, we do not yet have quantitative data to indicate whether this speculative generalization holds. Whatever the factors that cause various NPs to shrink from canonical subject position in copular constructions, the data show that while proportional and definite NPs strongly favor canonical subject position over pivot position, there are also NPs which strongly favor existential pivot position.

*Result 4: Sub-orderings among definites*

The strong NPs vary in existentiality by a factor of over 8000, and there is systematicity to the distribution. We treated first or second person (i.e. local) pronouns as a separate NP type from third person pronouns, in order to see whether there is any differential effect in their existentiality, although both types are already known to vastly prefer canonical over existential pivot position. What we found is that local pronouns are less existential than non-local pronouns, the expected result given that cross-linguistically local pronouns are prototypical subjects. Demonstratives were found to be more existential than regular pronouns, also in line with their being less canonical subjects, and distal demonstratives are more existential than proximal demonstratives, which is once again the expected result.<sup>9</sup>

*Result 5: Exclusives contribute to existentiality*

It is known that NPs headed by exclusives such as *only* occur in existential pivot position, but not known whether this is to be attributed to the effect of *only* or to the independent licensing in existential pivot position of the argument of *only* — see Beaver (2004). To check whether the exclusive *only* affects licensing, we included an NP form consisting of *only* followed by a local pronoun. What we find is that the presence of *only* makes the NP substantially more existential, by a factor of approximately 10,000. With the caveat that there are probably separate effects whereby unstressed NPs are more existential than stressed NPs, and complex NPs more existential than simple NPs, we conclude that *only* has an effect in making NPs more existential.

*Result 6: Genre strongly affects use of existentials*

Our studies using structured corpora are still in progress, but preliminary results are shown in figure 2. Due to the smaller size of the structured corpora relative to the web corpus, we were limited in the type of expressions we could meaningfully investigate. None the less, our data covers almost the same range of categories as the web based study.<sup>10</sup>

The broad trends in the web survey are repeated for all three structured corpora: pronouns are the least existential, downward monotone NPs are the most existential and show a clear anti-definiteness effect, and definites and prototypical indefinites fall in between with definites having considerably higher C/E ratios than indefinites. The broad similarity between the web results and the structured corpus

results provides evidence that despite the clear methodological problems involved in using an internet search engine for gathering quantitative data, the web results are in fact linguistically significant.

There is a wealth of further information that can be drawn out of figure 2, but for reasons of space we limit ourselves to just one further observation. Consider what is probably the most visually salient property of the figure: the WSJ graph is uniformly higher than the other two. In other words, writers (and editors) of newspaper text uniformly avoid existential constructions relative to what is found in the broad mix of genres in the Brown corpus, and relative to what is found in the type of everyday speech contained in Switchboard. For example, in newspaper text the prototypical indefinite *a(n)* is about three times as likely to occur in canonical position as in existential pivot position. For a mixed corpus, *a(n)* is slightly more likely to appear in existential pivot position (C/E ratio = 0.9), and for everyday speech *a(n)* is 2.5 times as likely to appear in existential position as canonical subject position. This is a striking example of genre variation which, to our knowledge, has not previously been noticed. We speculate that the distribution of indefinites might relate in part to the desire of newspaper writers to appear authoritative about the information they present, an effect which may be obtained if indefinites are given a specific reading. If this reading is more easily obtained in canonical subject position, pivot position will be disfavored for indefinites. However, the effect runs across the board of NP types (though it is particularly strong for indefinites), so there must be other factors, e.g. a desire to keep length down, and so avoid expletives that from an editor's point of view might appear unnecessary.

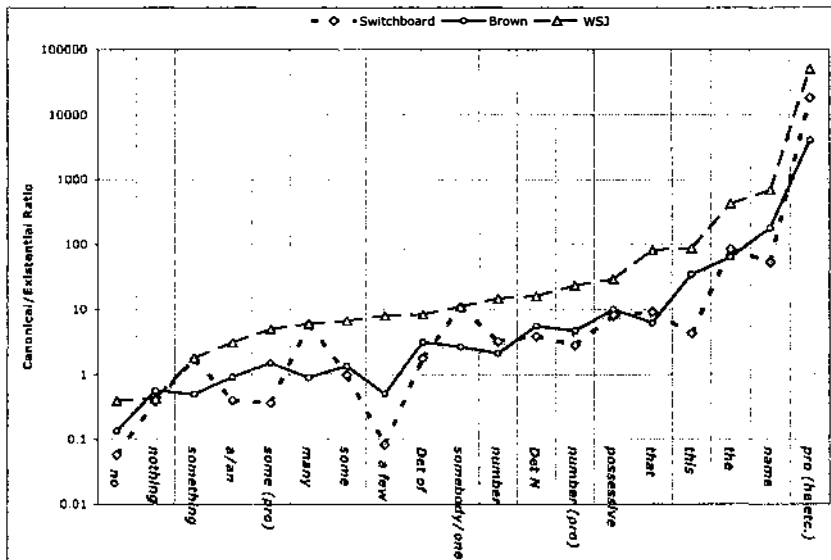


Figure 2: Comparison of C/E ratios by Genre

*Result 7: Dutch NP distribution globally mirrors English*

The Dutch existential construction has a similar form to the English existential, with an initial expletive *er* ‘there’ followed by verb and then non-canonical subject. The range of verbs that can occur in this construction is much wider in Dutch than in English, although, as for English, occurrences of the standard copular *zijn* ‘to be’ are prototypical. We performed a study of Dutch existentials using Google, examining only cases with the standard copular as main verb. The results are presented in the chart in figure 3. At the time we conducted the study (mid-2004) the amount of Dutch in the Google database was about 5% of the amount of English. Thus we were more constrained for Dutch than for English in our ability to conduct searches for infrequent patterns, and we did not attempt to fully replicate the range of English patterns we searched for.<sup>11</sup>

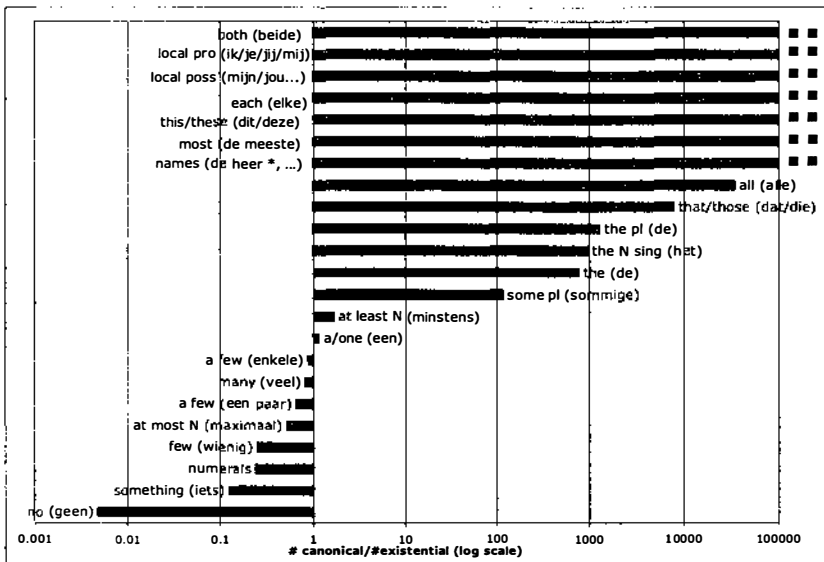


Figure 3: Dutch Canonical/Existential ratios

The results for Dutch are broadly similar to those for English. First, there is a clear definiteness effect consisting of a large gap in frequency space between NPs that would standardly be classified as strong, and those that would be classified as weak, and within the strong NPs. Second, within the strong NPs, definite NPs are considerably more existential than pronouns and NPs headed by proportional determiners. Third, prototypical indefinites such as *een* ‘a’ (which is not distinguished in our data from the stressed non-schwa variant *één* ‘one’) are far from being prototypical as existential NPs, and have a C/E ratio close to parity. The prototypically existential NPs include those headed by downward monotone determiners as well as those headed by numerals, and all of these exhibit a clear anti-definiteness effect.

*Result 8: Dutch Sommige ‘some’ behaves as if strong*

It is interesting to note that the Dutch indefinite *sommige* ‘some’ (here considered as a determiner, though it can also occur as a full NP) should clearly be classified as a strong NP. This is anomalous from the point of view of standard semantic theories of the strong/weak distinction. However, the results we obtained mirror the analysis of de Jong (1983) and de Hoop (1995) who show that *sommige* ‘some’ behaves as a strong determiner. We take this confluence of conclusions as providing support for our methodology.

*3.3. Discussion: The Order of NPs*

Since we found both definiteness effects and anti-definiteness effects, we may say that our quantitative studies of English and Dutch show the following: *the definiteness effect is a two sided coin*. We can summarize the main results in terms of the following scale:

- (27) Pronouns, proportional NPs > definite descriptions > prototypical indefinites > ↓mon NPs

By *prototypical indefinites* we mean whichever indefinites are most common in a language, though for languages other than those reported in this section there is a separate issue of whether indefiniteness is standardly marked in that language at all. Thus Dutch *sommige* ‘some’ is not a prototypical indefinite since it has much lower frequency than *een* ‘one’, and English numerals are also not prototypical, having much lower frequency than *a(n)* and less uses (e.g. no generic use).

Some further generalizations can also be seen in terms of scales. As regards pronouns, for example, the ordering we obtained is what would be expected on the basis of an analysis based on the subject property account we described: local pro > non-local pro. Similarly, for most of the studies we performed we obtain the natural ordering among demonstratives proximal > distal. However, it is interesting to note that for the Switchboard corpus this ordering was reversed: this is presumably because of a prominent use of English proximal demonstratives to present new discourse referents: “*there’s this guy I know...*” Finally, for the study based on the largest corpus (i.e. the English web study) we have pro > proportional NP, although for other studies we lacked sufficient data to substantiate this observation. Ignoring the subordering of demonstratives, we suggest the following more fine grained scale than that in (27):

- (28) Local pro > non-local pro > proportional NPs > definite descriptions, demonstratives > prototypical indefinites > ↓mon NPs

#### 4. Identifying Definiteness Effects: Two Case Studies from Russian

We now discuss two constructions which have been a major concern in work on Russian syntax. The first, Subject Genitive of Negation, is typically not seen as manifesting a definiteness effect. We show that this construction does exhibit definiteness effects, although the permissiveness of the construction relative to e.g. the English existential means that the we can observe the effects not by looking at which NPs are forbidden, but at which NPs are preferred. That is, what we actually see are anti-definiteness effects. The second construction is the optional copula in present tense locatives/possessives, which has been claimed elsewhere to show the definiteness effect. Using quantitative data, we show that the effect in this construction is somewhat different from the definiteness effects observed in the other constructions we have considered.

##### 4.1. Subject Genitive of Negation: An Anti-definiteness Effect

The choice of the canonical or the existential form for negative sentences in Russian has been discussed due to the case marking of the NPs in such sentences. While the subject of the canonical sentences has the Nominative case, the subject of the negated existential has the Genitive case (this case marking, known as Genitive of Negation, is optional for direct objects in negated sentences). In the past tense, the copula in the canonical sentences is in agreement with the subject (1), while the copula in the existential sentences (2) is always in neuter singular (examples from (Partee and Borschev, 2002)).

- (1) Nikto tam ne byl.  
No one.NOM there neg was.M.SG.  
'No one was there.'
- (2) Nikogo tam ne bylo.  
no one.GEN there neg was.N.SG  
'No one was there.'

The factors responsible for the choice of the form and the meaning differences between the forms have been much debated. The factors proposed to distinguish between the forms are Topic/Focus structure (Babby, 1980), agentivity (Padučeva, 1992), and perspectival structure (Partee and Borschev 2001; 2002; 2004). It is clear from the latter publications that a purely Topic/Focus explanation makes incorrect predictions, and that both agentivity and perspectival structure play a role in determining the choice of construction.

At first, the choice of the canonical or the existential construction in these cases does not seem connected to the definiteness effect, as definite and indefinite phrases alike occur freely in the existential construction. We checked the frequencies of the variants for some noun phrases in past tense sentences<sup>12,13</sup>. Table 2 contains the frequency of canonical and existential constructions with the locative phrase *tam* 'there' for the word order NP LocP V, as in (29) and (30).

- (29) Èta kniga tam ne byla.  
This book.NOM there not was.F  
'This book wasn't there.'
- (30) Ètoj knigi tam ne bylo.  
This book.GEN there not was.NEUT.  
'There wasn't this book there.'

NP	translation	canon.	exist.	exist./total
<i>ja</i>	I	11,700	6,100	34%
<i>moja</i> *	my * (f.)	3	100	97%
<i>èta</i> *	this * (f.)	0	100	100%
<i>èti</i> *	these *	2	80	97%
<i>takaja</i> *	such * (f.)	0	80	100%
<i>takie</i> *	such * (pl.)	0	90	100%

Table 2: Canonical and existential constructions in negative past tense sentences of the form *NP tam ne V*.

As the table shows, most NPs rarely occur in the canonical form. This form is only used substantially with animates, which is reasonable, given that this form signals agentivity. Table 2 shows a strong anti-definiteness effect, with most NPs occurring almost exclusively in the existential constructions. The subject properties requirements for the canonical form in this case are high, and are only satisfied by NPs high on the prominence scales. We can summarize the data from the table in terms of the following NP ordering:

- (31) local pro > local possessive > demonstratives, indefinites

The ordering in (31) is a sub-ordering of that observed in our English data. Therefore, we suggest that despite the enormous distributional differences between pivot position in English existentials and the Russian Genitive of Negation, both distributions be seen as instances of a definiteness effect, and both distributions be seen as involving an NP occurring in a non-canonical construction because it lacks canonical subject properties.

#### 4.2. *The Optional Copula in Present Tense Possessives/Locatives: A Definiteness Effect?*

In this section we consider a Russian construction, namely possessives/locatives in the present tense, that has been claimed to show a definiteness effect. We use quantitative data to show that the effect departs significantly from the canonical/existential alternations discussed above.

In Russian, positive locative and possessive sentences in present tense can appear with or without the copula (Kondrašova, 1996, Chapter 5.2). The variant with the copula typically has a non-exhaustive meaning, and the variant without has an exhaustive meaning. For example, the English sentence (32) can have two meanings: *Some of John's teachers are good* or *(All) John's teachers are good*. The former sense is expressed by the Russian sentence (33), with the copula, and the latter sense is expressed by sentence (34), without the copula.

- (32) John has good teachers.
- (33) U Džona jest' xorošie učitel'ja.  
at John *jest'* good teachers.  
'Some of John's teachers are good.'
- (34) U Džona xorošie učitel'ja.  
at John good teachers.  
'(All) John's teachers are good.'

Kondrašova (1996, 2.4.4) claims that the constructions with the copula *jest'* show the definiteness effect. In our terminology, the variant with the copula is the existential, and the one without is its canonical alternative. Her examples and judgments are:

Strong quantifiers vs. weak quantifiers:

- (35) \*V Moskve jest' malo milicionerov / vse lingvisty.  
in Moscow *jest'* few policemen.GEN / all linguists.NOM  
'There are few policemen/all the linguists in Moscow.' (= her (40))
- (36) V Moskve jest' nemalo milicionerov / neskol'ko parkov.  
in Moscow *jest'* not.few policemen.GEN / several parks.GEN  
'There are quite a few policemen/several parks in Moscow.' (= her (41))

Definite vs. Indefinite descriptions:

- (37) \*V Moskve jest' tot čelovek, kotoryj kupil kartinu iz Èrmitaža.  
In Moscow *jest'* that person who bought painting from Hermitage  
'In Moscow there is the person who bought a painting from the Hermitage.' (= her (42a))
- (38) V Moskve jest' takie ljudi, kotorye pokupajut kartiny iz Èrmitaža.  
In Moscow *jest'* such people who buy paintings from Hermitage  
'In Moscow there are people who buy paintings from the Hermitage.' (=her (43a))

There are two problems with Kondrašova's description. First, according to the classification of Milsark (1977), *malo* is not a strong NP. Strong NPs involve quantification over a domain, while weak ones are number determiners or "cardinality words". According to this classification, *malo* in its cardinal interpretation is a weak NP, so it should be felicitous in an existential. The same prediction would be made by Keenan (2003), as *malo* is what he terms "conservative on the second argument." Nevertheless, *malo* is not used with the copula even in the cardinal interpretation.

Second, the searches we made showed that although the existential ratio of *malo* 'few' vs. *nemnogo* 'some' and *mnogo* 'much' is indeed substantial, the difference between *tot* \* *kotoryj* 'the \* that' (m.) and *takoj* \* *kotoryj* 'a \* that' (m.) is not as large, and there are many examples with and without the copula for each of these constructions — see figure 4. This is different from the effect observed in English, as the difference in the ratio between the definite and indefinite NPs in English is much higher. What is more striking, the word *malo* has a much higher C/E ratio than the indefinites, while the word 'few' in English has a ratio similar to other indefinites. Although more extreme, the unexpected behavior of *malo* is reminiscent of the Dutch *sommige* 'some' (see section 3), which also behaves as a strong NP despite denoting a cardinal quantifier. Further investigation of the semantics of *malo* is required in order to understand this behavior.

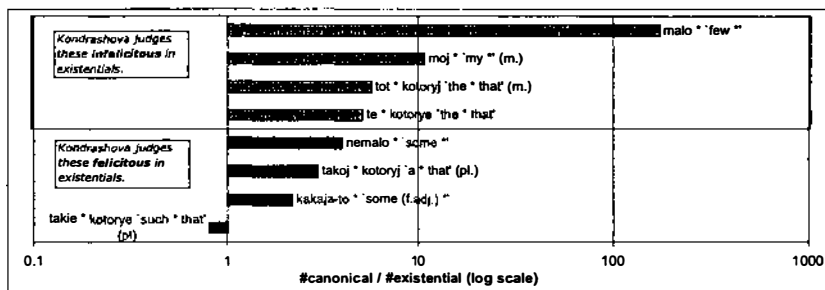


Figure 4: Usage of copula in Russian present tense locatives/possessives.

The difference between the ratios of the quantifiers in the table, except *malo*, is not large, and is not enough to establish whether the definiteness effect exists as in the constructions discussed in section 3. The theoretical reason for this may be that, in this case, presence or absence of *jest'* has only a minor impact on the subjecthood of the NP. When the copula is present, the NP is farther from its canonical location in the beginning of the sentence than when the copula is absent. The form and agreement properties of the NP are the same. This may be the reason for the smaller differences in the canonical/existential ratio.

## 5. Conclusion

We have revealed a complex distribution of NP types in existential constructions involving both definiteness and anti-definiteness effects. We have shown, using quantitative data, that the definiteness effect is not categorically predictable from NP form. The tendency of different kinds of NPs to appear in existential, as opposed to canonical, constructions, varies widely, but there are many possible levels between the extremes. The explanation for the distributional properties of the NPs is given in terms of their subject properties. An NP can be realized as an existential pivot when it lacks properties that would make it a good subject, so NPs that possess more subject-like properties tend to occur more frequently in canonical constructions.

We have argued that this analysis makes correct predictions for a number of distributional issues in Dutch, English, Hebrew and Russian, and we have discussed some challenging examples of cross-linguistic variation, cases in which the range of existential pivot NP types is different in one language from that in another. However, when the distribution depends on presence/absence of subject properties, it is systematically linked to underlying orderings of NPs, such as:

Local pro > non-local pro > proportional NPs > definite descriptions,  
demonstratives > prototypical indefinites > ↓mon NPs;

NPs with token interpretations > NPs with type interpretations.

If we are right, then the orderings we have found should be universal: in any construction involving non-canonical realization of a subject, and which is in competition with a canonical subject construction, the directionality of this ordering will be preserved. This hypothesis is suggestive of future cross-linguistic research: more languages, and more constructions. We believe that our quantitative, corpus based methodology based on the use of canonical/existential ratios can be applied fruitfully to other languages, and may be extended to other constructions for which a competition-based analysis is plausible.

## Endnotes

\* This paper is based on research conducted by a joint Stanford/UC Berkeley research group run by David Beaver and Line Mikkelsen; additional members are Gerlof Bouma, Alex Bratkievich, Iván García Álvarez, Florian Jaeger and Laura Whitton. We benefited from extensive discussions with Gregory Ward and from comments of audiences at New York University, Northwestern University and Stanford University, the KNAW Academy Colloquium *Cognitive Foundations of Interpretation*, and SALT XV. Special thanks to Judith Tonhauser for valuable comments on the paper itself.

<sup>1</sup>Throughout this paper we use ‘NP’ to refer to all nominal phrases, circumventing the NP/DP distinction, which is irrelevant to the issues discussed here.

<sup>2</sup>Lambrech (2000) discusses the object-like properties of pivots cross-linguistically.

<sup>3</sup>See e.g. Woolford (2005), who shows that local pronouns are less marked than non-local pronouns for case marking purposes.

<sup>4</sup>The structured corpus searches used the utility *tgrep2* on Penn Treebank corpora pre-parsed using the Treebank 2 annotation guidelines: Godfrey et al. (1992), Marcus et al. (1994).

<sup>5</sup>The searches were carried out over a period of a few days in August 2004. Both the web and search engine technology evolve constantly, so repeating the tests now would yield numerically distinct results. We would expect the ratios we calculate to remain roughly constant provided the makeup of the web involves roughly constant proportions of given genres, and provided the time period is small enough to ignore language change. A problematic issue is that the functionality of Google's search engine has changed several times in the last year; at time of writing, the behavior of the wildcard operator \* is not clear. Other search engines (e.g. Yahoo) also allow wildcards, but the functionality of these wildcards is also unclear.

<sup>6</sup>Search engines order their results in a way that introduces bias as regards the types of text which tend to be highly ranked in a search. Therefore it would be advisable to sample a random selection of search results rather than e.g. the first 20. However, for the English Google data we sampled only the first few pages of search results.

<sup>7</sup>We experimented with additional searches that combine the regular search with a negative criterion using Google's "-" feature to block unwanted results like *a few*. However, this introduces complications, first because Google's count results may not be consistent across searches with and without negative search criteria, and second because such searches block legitimate search results in case the search pages contain an instance of the blocked expression elsewhere. Therefore we did not use negative criteria in any of our searches, and instead relied solely on sampling to pick up on erroneous data.

<sup>8</sup>Numerals are the one class of weak determiners for which our web results were not duplicated in structured corpora, so their high existentiality in our web data may result from an anomaly in our search procedures.

<sup>9</sup>We did not originally make a separate search for *it*, but a search at time of writing indicates that *it* is even less existential than local pronouns. This effect runs against the grain of universal hierarchies in which third person is less prototypical as a subject than first and second, and inanimate NPs are less prototypical than animate NPs. However, *it* is an intrinsically weak pronoun (it's strong counterpart, arguably, being *this*), so when we compare *it* with other pronouns there is a confound: we are comparing weak occurrences of one pronoun with both weak and strong occurrences of others.

<sup>10</sup>In figure 2, NPs are ordered according to their C/E ratio in the WSJ corpus. For this reason it appears in the graph that in some sense NPs in the WSJ corpus have a smooth monotonic distribution while the distribution of NPs in the other corpora is more jumpy. This is an illusion created by the choice of NP ordering.

<sup>11</sup>The Dutch searches were sampled in the same way as the English searches. However, the person who performed the sampling (Beaver) is not a native speaker of Dutch. Ideally the searches would be repeated using a native speaker. An additional

complication is that in Dutch the criteria for determining whether an occurrence of *er* 'there' is an expletive or a locative is even more vexed than for English. For example, occurrences of *er* 'there' in immediately post-verbal position are more common in Dutch than in English and some of these occurrences may involve an expletive. Since we take non-canonical position (or other marking) of the subject as the crucial issue, we did not count such tokens with existentials.

<sup>12</sup>Some discussion of the frequencies and the meanings of the variants is found in Partee and Borschev (2002). The definiteness effect is not mentioned there.

<sup>13</sup>We checked the sentences in the negative past tense form for technical reasons: in this form the canonical/existential variants differ in a way that makes searching practical.

## References

- Aissen, Judith: 1999, "Markedness and Subject Choice in Optimality Theory", *Natural Language and Linguistic Theory* 17, 1–39.
- Babby, Leonard H.: 1980, *Existential Sentences and Negation in Russian*, Karoma Publishers, Ann Arbor.
- Barwise, Jon and Robin Cooper: 1981, 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* 4, 159–219.
- Beaver, David: 2004, 'Five 'Only' Pieces', *Theoretical Linguistics* 30, 45–64.
- Bentley, Delia: 2002, 'Definiteness Effects: Evidence from Sardinian', Ms., University of Manchester.
- Borschev, Vladimir and Barbara Partee: 2001, 'The Russian Genitive of Negation in Existential Sentences: The Role of Theme-Rheme Structure Reconsidered', In Hajičová, Eva, Petr Sgall, Jiří Hana, and Tomáš Hoskovec (eds.), *Travaux de Cercle Linguistique de Prague*, v. 4, 185–250, John Benjamins, Amsterdam.
- Clark, Eve: 1978, 'Locational: Existential, Locative, and Possessive Constructions', In Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik (eds.), *Universals of Human Language*, Vol. 4, 85 – 126, Stanford University Press, Stanford, CA.
- Freeze, Ray: 1992, 'Existentials and Other Locatives', *Language* 68, 553–95.
- Givon, Talmy: 1978, 'Definiteness and Referentiality', In Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik (eds.), *Universals of Human Language*, Vol. 4, 291–330, Stanford University Press, Stanford, CA.
- Godfrey, J., E. Holliman, and J. McDaniel: 1992, 'SWITCHBOARD: Telephone Speech Corpus for Research and Development', In *ICASSP92*, 517–520, IEEE, San Francisco.
- de Hoop, Helen: 1995, 'On the Characterization of the Weak-Strong Distinction', In Bach, E., E. Jelinek, A. Kratzer, and B.H. Partee (eds.), *Quantification in Natural Languages*, 421–450, Kluwer, Dordrecht.
- de Jong, Francisca: 1983, 'Sommige Niet, Andere Wel; de Verklaring van een Raadselachtig Verschil', *GLOT* 6, 229–246.
- Keenan, Edward: 1976, 'Towards a Universal Definition of "Subject"', In Li, Charles (ed.), *Subject and Topic*, Academic Press, New York.

- Keenan, Edward: 2003, 'The Definiteness Effect: Semantics or Pragmatics?', *Natural Language Semantics* 11, 187–216.
- Kondrašova, Natalia Yurievna: 1996, *The Syntax of Existential Quantification*, Ph.D. thesis, University of Wisconsin, Madison.
- Lambrecht, Knud: 2000, 'When Subjects Behave Like Objects: An Analysis of the Merging of S and O in Sentence Focus Constructions Across Languages', *Studies in Language* 24, 611–682.
- Marcus, Mitchell P., Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger: 1994, 'The Penn Treebank: Annotating Predicate Argument Structure', In *ARPA Human Language Technology Workshop*, 114–119, Morgan Kaufmann, Plainsboro, NJ.
- McNally, Louise: 1998, 'Existential Sentences with Existential Quantification', *Linguistics and Philosophy* 21, 353–392.
- Mikkelsen, Line: 2002, 'Reanalyzing the Definiteness Effect: Evidence from Danish', In *Working Papers in Scandinavian Syntax*, volume 69.
- Milsark, Gary: 1977, 'Toward an Explanation of Certain Peculiarities of the Existential Construction in English', *Linguistic Analysis* 3, 1–29.
- Moro, Andrea: 1994, *The Raising of Predicates*, Cambridge University Press.
- Padučeva, Elena: 1992, 'O Semantičeskome Podxode k Sintaksisu i Genitivnom Sub'ekte Glagola byt', *Russian Linguistics* 14, 53–63.
- Partee, Barbara H. and Vladimir Borschev: 2002, 'Genitive of Negation and Scope of Negation in Russian Existential Sentences', In *Formal Approaches to Slavic Linguistics* 10.
- Partee, Barbara H. and Vladimir Borschev: 2004, 'The Semantics of Russian Genitive of Negation: The Nature and Role of Perspectival Structure', *Proceedings of Semantics and Linguistic Theory* 14.
- Reinhart, Tanya: 2005, 'Topics and the Conceptual Interface', In Kamp, Hans and Barbara Partee (eds.), *Context*, Elsevier.
- Silverstein, Michael: 1976, 'Hierarchy of Features and Ergativity', In Dixon, Richard (ed.), *Grammatical Categories in Australian Languages*, Australian Institute of Aboriginal Studies.
- Ward, Gregory and Betty Birner: 1995, 'Definiteness and the English Existential', *Language* 17, 722–742.
- Woolford, Ellen: 2005, 'Differential Subject Marking at Argument Structure, Syntax, and PF', Ms., University of Massachusetts. To appear in Helen de Hoop and Peter de Swart eds., *Differential Subject Marking*. SNLLT series, Kluwer Academic Publishers.
- Ziv, Yael: 1982, 'Another Look at Definites in Existentials', *Linguistics* 18, 73 – 88.
- Zucchi, Alessandro: 1995, 'The Ingredients of Definiteness and the Definiteness Effect', *Natural Language Semantics* 3, 33–78.