

Gradual Uniqueness Effect in Non-Unique Definite Noun Phrases

Dmitry Levinson
Stanford University

January 7, 2006

1 Uniqueness effect

1.1 Examples

Using the definite article is usually accompanied by the *uniqueness effect*. This term describes the implication that the referent of the noun phrase is unique or uniquely identifiable, that there is only one most salient candidate for being the referent. Examples:

Sentence (1) suggests that there is exactly one clown in the puzzle.

- (1) Teacher, giving directions: On the next page, you will find a puzzle. Find the clown in the puzzle. [example (3) in (Roberts 2003)]

The following sentence shows that only the definite article is appropriate if the description is semantically unique:

- (2) Last weekend we climbed the/[#]a biggest mountain in West Virginia. [example (11) in (Roberts 2003)]

The following question suggests that there was only one inventor of calculus:

- (3) Who is the inventor of the calculus? [= (8) in Szabó (2000)]

1.2 Explanations

Two kinds of explanations:

- Semantic (Kadmon 1990, Roberts 2003). Uniqueness – part of the conditions for using the definite article. Uniqueness is part of the semantic component of the definite article and its presence or absence determines a noun phrase's being definite or indefinite.
- Pragmatic (Heim 1982, Szabó 2000). The uniqueness – consequence of the definite article usage. Sometimes – there is no uniqueness effect.

1.3 Body parts – a counterexample?

Body part terms are frequently used with the definite article even when there is no uniqueness.

- (4) Juan se torció *el pie*. [from Kliffer (1983)]
 (5) Moví *la mano*. [from Kliffer (1983)]
 (6) The dog bit him on *the finger*. [Epstein (1999)]
 (7) I broke *my toe*. [Epstein (1999)]

A person usually has two hands and two legs, therefore, if the relevant body part was not made salient in the previous discourse, these examples show a use of definite article without the uniqueness effect.

Such usage is also frequent with other *relational* nouns:

(8) I hope the café is located at *the corner of a busy intersection*. [Barker (in press)]

(9) He was *the son of a poor farmer*. [Löbner (1985)]

1.4 Common assumption

(10) The uniqueness is *categorical*. That is, the only thing that matters is whether the number of the candidates is one or more.

Therefore, examples with two equally salient candidates are treated the same as examples with a large number of candidates.

2 Article usage with body part terms: data

2.1 Cardinality

Birner (1988:138) noticed that in the variation between the possessive and indefinite in English,

...there seems to be a strong correlation between the number of relevant body parts and the relative acceptability of the possessive pronoun vs. the indefinite article. Very consistently, the fewer the body parts involved, the more acceptable the use of the possessive and the less acceptable the use of the indefinite article.

Her examples are:

(11) a. I bumped my/[#]a nose.

b. I broke my/[?]an arm.

c. I broke my/a finger.

d. I tore [?]my/a ligament.

e. I burst [#]my/a blood vessel.

I check the usage of determiners with different body part terms in a number of languages. Table 1 lists some body part terms and their cardinality in the human body.

Name	Cardinality
nose, head	1
hand, arm, leg, foot, eye, ear	2
finger	10
tooth	up to 32

Table 1: Cardinality of Body Parts.

2.2 Sentences

(12) *Me duele la/[?]una mano*. SPANISH

Me hurts the/a hand.

'My hand hurts.'

(13) *Se rompió un/el brazo*. SPANISH

'He/she broke his/her arm'.

(14) *I broke a/^{*}the/my finger*.

2.3 Native Speaker Intuitions - Spanish

For unique body parts, only the definite article can be used (15). For dual body parts, the definite article is preferred (16). For fingers, both the definite and the indefinite article can be chosen (17). Interestingly, if the finger is named so that there are only two body parts with that name, the definite article is preferred (18). For teeth, the definite article suggests that the tooth that hurts was salient in some way (19). These intuitions are similar to those reported by Birner (1988) for English.

(15) *Me duele la/*una cabeza/nariz.*

Me hurts the/a head/nose.
'My head/nose hurts.'

(16) *Me duele la[#]una mano.*

Me hurts the/a hand.
'My hand hurts.'

(17) *Me duele el / un dedo.*

Me hurts the / a finger.
'My finger hurts.'

(18) *Me duele el / [?]un dedo índice.*

Me hurts the / a finger index.
'My index finger hurts.'

(19) *Me duele [#]el / un diente.*

Me hurts the/ a tooth.
'My tooth hurts.'

2.4 WWW results

I examined the distribution of articles in sentences like (12) - (14), with different body part terms. I performed the search on the World Wide Web, using Google search engine. The results, in Table 2 and Figure 1, show considerable variation in the article choice among body part terms:

Body part	Spanish: <i>duele</i>	Spanish: <i>se rompió</i>	English: <i>I broke</i>
<i>hand</i>	2% [5/211]	37% [26/69]	1% [13/959]
<i>arm</i>	3% [6/178]	48% [103/214]	2% [114/5700]
<i>leg</i>	12% [20/172]	61% [273/444]	4% [300/7300]
<i>foot</i>	16% [13/83]	---	1% [20/2300]
<i>ear</i>	17% [29/167]	---	---
<i>eye</i>	19% [27/139]	---	---
<i>finger</i>	18% [18/98]	65% [70/107]	29% [280/940]
<i>tooth</i>	61% [47/77]	91% [39/43]	73% [901/1300]

Table 2: Percentage of indefinite article usage with non-unique body part terms. [In brackets: the number of occurrences with the indefinite article/the total number of occurrences].

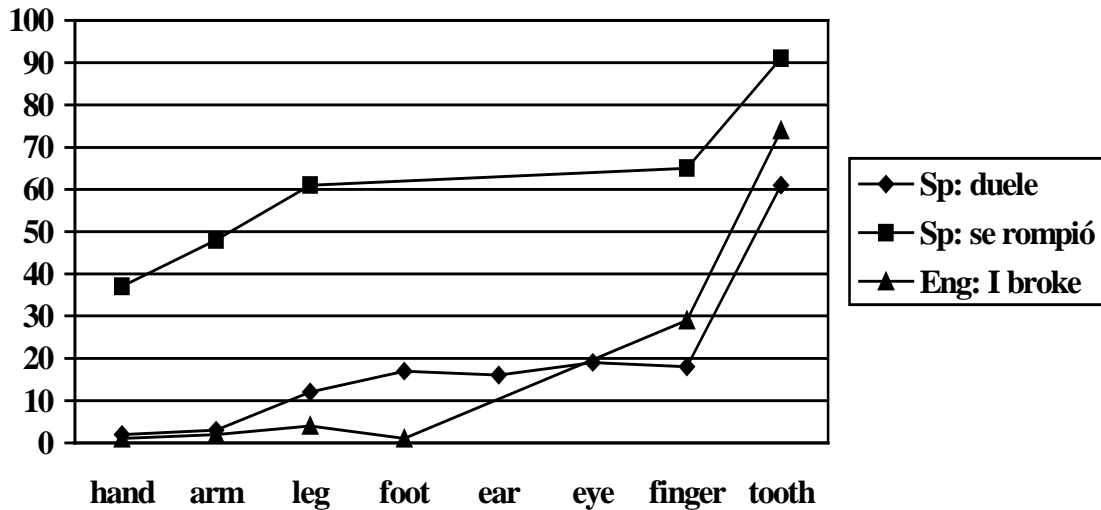


Figure 1. Percentage of indefinite article usage with different body part terms.

For unique body parts, only the definite determiner is allowed. For non-unique body parts, there is a variation between the definite and the indefinite determiner. The body part *cardinality* was shown to have a statistically significant correlation with the percentage of the indefinite article usage.

Similar effects were observed for Italian and Brazilian Portuguese (Levinson 2005).

2.5 Syntactic and lexical context

Birner (1988) noticed that the syntactic construction in which the noun phrase is used is also a factor affecting the choice of the determiner, although this is unexpected on the basis of previous accounts. For example, the indefinite article is not used at all in prepositional constructions like (20). Moreover, corpus data show that within the same syntactic construction, the choice of the article in the noun phrase may depend on other elements of the sentence. For example, when the NP is the direct object, the article usage depends on the verb (Figure 2).

(20) *I hit him on the/*an arm.*

2.6 Conclusions

Two kinds of factors affecting the choice of determiner:

- The noun: cardinality.
- The syntactic and lexical context.

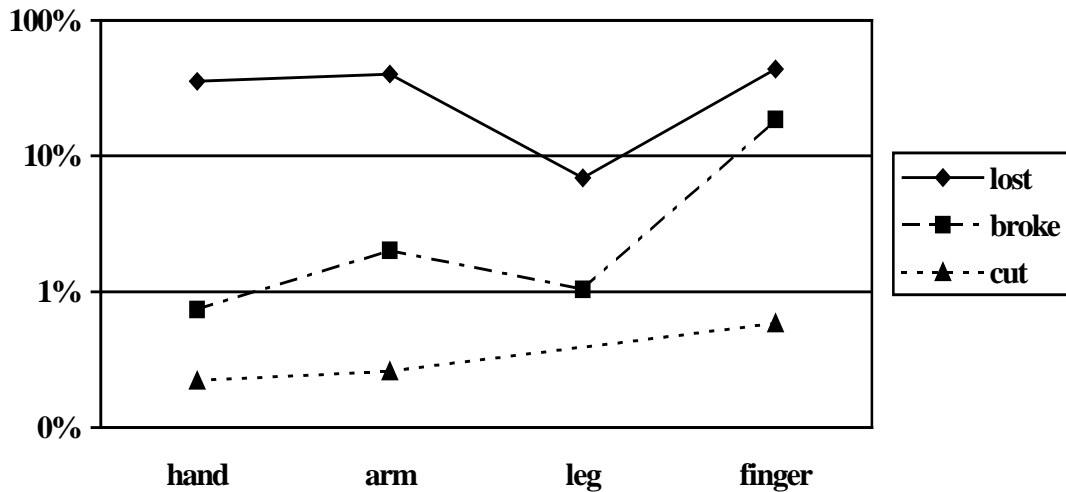


Figure 2. Percentage of indefinite article usage in sentences of the type ‘I V my/a N’ with different verbs (Google search).

3 Analysis

3.1 Identifiability, cardinality and the context

The combination of various factors governing the choice of the article results in *gradual transition* between preferring the definite article and preferring the indefinite one. This supports the view that the conditions for the article use are not categorical. The choice of the article is affected by the *identifiability* of the body part, which is affected by its cardinality. For relational nouns identifiability doesn’t have to be complete in order for the definite/possessive determiner to be felicitous. Instead identifiability is a constraint that can be violated to different degrees when the definite/possessive is used with non-unique NPs. Lower identifiability leads to more occurrences of the indefinite article. If the identifiability is high, the indefinite article is less probable.

higher *cardinality* \Rightarrow lower identifiability \Rightarrow the indefinite article is more frequent

What is then the role of the context? Du Bois (1980:232) suggests that the identifiability is required to a reasonable extent. The *required precision* may be different for different syntactic and lexical contexts. Higher required precision means that higher levels of identifiability are required for the same usage of the definite article. With constant identifiability, a higher required precision leads to a smaller probability of the definite article.

higher *required precision* \Rightarrow the indefinite article is more frequent

How does the required precision depend on the context? A reasonable hypothesis would be that the more prominent and affected is the noun phrase, the higher the required precision.

For the syntactic construction, the prediction is that the indefinite article would be more frequent with main arguments (direct objects) and less frequent with indirect objects and adjuncts. This explains the difference between (14), in which the body part is the direct object, and (20), in which it is less prominent.

For the lexical context, when the NP is the direct object, the indefinite article would be more frequent with verbs that denote change of state that with those that do not. Among verbs denoting change of state, the indefinite article would be less frequent with verbs denoting smaller change of state. This was confirmed in a preliminary investigation (Figure 2).

3.2 Formalization

The analysis proposed here is based on the approach to definiteness described in Levinson (2000:63). In this analysis, the definiteness and the indefiniteness markers form a scale <DEF, INDEF>. The definiteness marker is a sign that the identifiability of the referent is at an appropriate level. According to Grice's (1975) maxim of Quality, it should not be used when the identifiability is less than required. The indefiniteness marker does not signal identifiability. It is, however, accompanied by a Q-type implicature (Levinson 2000:41) that the identifiability was not high enough for the definite marker. Such an implicature happens due to the maxim of Quantity.

The way these maxima influence the actual usage can be modeled within a theory of violable constraints of different ranks, allowing for degrees of violation. The predicted probability of a variant is inversely proportional to a monotone function of its combined degree of constraint violations, which is the weighted sum of the ranks of the violated constraints, with degrees of violation as weights.

$$total_violation(i) = \sum_j r_j \cdot d_{i,j}$$

r_j - rank of constraint j ; $d_{i,j}$ - degree of violation of constraint j for variant i

The constraints are listed below:

- *DEF. This constraint represents the maxim of Quality. I propose that for the relational nouns the rank is *ReqP* (*required precision*), varying with the syntactic construction. The degree of violation is (*cardinality* - 1). This constraint penalizes for signaling more identifiability than present.
- *INDEF. This constraint represents the maxim of Quantity and the need to avoid unnecessary Q-type implicatures. For our purposes, the rank and the degree of violation can be constant and equal to 1.

The constraints are summarized in the following table:

Constraint	Rank	Degree of violation
*INDEF	1 (constant)	1 (constant)
*DEF	<i>ReqP</i>	(<i>cardinality</i> - 1)

Table 3. The constraints

The competition table for a general case is shown in Table 4. The higher the cardinality of the possible referents set, the higher the total violation for the *def* variant, and the more probable is the *indef* variant. This explains the gradual uniqueness effect shown in the data above.

	*indef [rank: 1]	*def [rank: <i>ReqP</i>]	total
indef	*: degree = 1	-	1
def	-	*: degree = cardinality-1	<i>ReqP</i> *(cardinality-1)

Table 4. Competition table for a general case

Special cases:

1. When the intended referent is unique, and full identifiability has been achieved, the *def* variant doesn't violate any constraint, and it is predicted to appear in all the cases:

	*indef [rank: 1]	*def [rank: <i>ReqP</i>]	total
indef	*: degree = 1	-	1
def	-	*: degree = 0	0

Table 5: Special case: full identifiability

2. In some syntactic constructions *ReqP* is very low, and the *def* variant is preferred, regardless of the cardinality of the possible referents set:

	*indef [rank: 1]	*def [rank: 0]	total
indef	*: degree = 1	-	1
def	-	*: degree = cardinality-1	0

Table 6: Special case: a syntactic construction with zero or very low *ReqP*.

3. The presented analysis is not inconsistent with the notion of the categorical uniqueness. If *ReqP* is very large, the *def* variant is predicted for *cardinality* = 1, and the *indef* variant is predicted for *cardinality* > 1. Therefore, categorical uniqueness is a special case of the analysis presented in this paper.

	*indef [rank: 1]	*def [rank: <i>ReqP</i> - very high]	total
indef	*: degree = 1	-	1
def	-	*: degree = cardinality-1	0 if cardinality = 1; very high otherwise

Table 7. Special case: Very high *ReqP* - categorical uniqueness.

References

- Abbott, Barbara (2004). Definiteness and Indefiniteness. In: Laurence R. Horn and Gregory Ward (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell.
- Barker, Chris (in press). Possessive Weak Definites. In: Kim, Ji-yung, Lander, Yury, and Partee, Barbara H. (eds). *Possessives and Beyond: Semantics and Syntax*. Amherst, MA.
- Birner, Betty J. (1988). Possessives vs. Indefinites: Pragmatic Inference and Determiner Choice in English. *IPrA Papers in Pragmatics* 2, 136-146.
- Birner, Betty J., and Gregory Ward (1998). *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins.
- Demonte, Violeta (1988). El “Artículo en Lugar del Posesivo” y el Control de los Sintagmas Nominales. *Nueva Revista de Filología Hispánica* 36, 89-108.
- Du Bois, John W. (1980). Beyond Definiteness: The Trace of Identity in Discourse. In: Wallace L. Chafe (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, 203-274. Norwood, NJ: Ablex Publishing Corporation.
- Grice, H. P. (1975). Logic and conversation. In: Cole & Morgan (eds), *Syntax and Semantics* 3.
- Heim, Irene (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. Dissertation, University of Massachusetts.
- Kadmon, Nirit (1990). Uniqueness. *Linguistics and Philosophy* 13, 273-324.
- Kliffner, M.D. (1983). Beyond Syntax: Spanish Inalienable Possession. *Linguistics* 21, 759-794.
- Levinson, Dmitry (2005). *Definiteness of Noun Phrases with Body Part Terms*. Manuscript.
- Levinson, Stephen C. (2000). *Presumptive Meanings*. MIT Press.
- Löbner, Sebastian (1985). Definites. *Journal of Semantics* 4, 279-326.
- Roberts, Craige (2003). Uniqueness in Definite Noun Phrases. *Linguistics and Philosophy* 26, 287-350.
- Szabó, Zoltán Gendler (2000). Descriptions and Uniqueness. *Philosophical Studies* 101, 29-75.