Computational Linguistics in the Netherlands Journal 1 (2011) 41-59 Submitted 7/2011; Published 12/2011

Using corpora tools to analyze gradable nouns in Dutch

Nick Ruiz NICHOLAS.RUIZ@GMAIL.COM
Edgar Weiffenbach EDGARWEIFFENBACH@GMAIL.COM
University of Groningen

Abstract

In this paper, we expand Morzycki (2009)'s claims that degree readings of size adjectives are attributed to syntax. We introduce a corpus-based analysis in Dutch to verify and extend his claim into the semantic domain. Using the LASSY Treebank, we extract syntactic and semantic properties of noun phrases consisting of the adjectives "gigantisch", "kolossaal", and "reusachtig" and manually annotate each adjective-noun pair with a gradable or non-gradable label.

Using these features, we construct a statistical model based on logistic regression and find that the grammatical role, definiteness, and particular semantic noun groups derived from Cornetto (a Dutch WordNet with referential relations) have a significant effect on the likelihood that an adjective-noun pair is interpreted by the reader to have a degree reading.

1. Introduction

Morzycki (2009) claims that degree readings of size adjectives, such as *a big idiot* are not merely the "consequence of some extragrammatical phenomenon," but rather can be attributed to syntax. He argues that "the syntax of a phrase gives rise to positional restrictions on the availability of these readings, and the semantics of degree measurement interacts with the scale structure of size adjectives to give rise to restrictions on the adjective itself." Morzycki (2009) focuses on gradable predicates provided by a noun and an adjective with a size reading.

1.1 Gradability

The adjectives in the phrases in (1) do not have a "physical size" reading, but rather denote a "high degree". In each of these cases, the noun has an inherent scale that can increase based on the attributive adjective.[1]

(1) a. een enorme idioot
an enormous idiot
b. gigantische consequenties
gigantic consequences

However, the nouns in (2) do not have an inherent scale, and thus can never be interpreted with a gradable reading.

1. For a discussion on why the inherent scale of gradable nouns can only increase, read Morzycki (2009)'s discussion of the "Bigness Generalization".

(2) a. Het kolossale tv-scherm
The colossal tv screen
b. Een reusachtige doos
A giant box

1.2 Research Hypotheses

In this paper, we look at how syntax and semantic grouping of nouns can contribute to the plausibility of a gradable reading. We focus on three size adjectives of similar meaning and etymology: “gigantisch” (gigantic), “reusachtig” (giant), and “kolossaal” (colossal). Using dependency structures, we retrieve contextual information about potentially gradable noun phrases to determine if the context can give us a hint at whether or not the NP is intended to be interpreted by the reader as a degree reading. We look at the following contextual features: *grammatical role*, *determiner type*, the presence of *adjunct prepositions* and *noun groups.*

Based on our intuitions we form three hypotheses. A first hypothesis is that NPs in the grammatical roles of object or predicative complements are more likely to allow a gradable reading than NPs in the subject position of the sentence. In addition, we hypothesize that NPs with indefinite readings are more prone to having a gradable reading than definite readings. NPs are indefinite if they have an article that does not indicate an identifiable object (i.e. “een” in Dutch, or “a” in English). Zero (or null) articles are typically considered in Dutch and English to be indefinite.

Another feature worth analyzing is the adjunct preposition. In (3), we see an example encountered in our corpus.

(3) een gigantische steun voor de nabestaanden van de slachtoffers
a huge support for the families of the victims

We hypothesize that the presence of an adjunct preposition has no effect on the plausibility of a gradable reading.

As mentioned earlier, we are also interested in exploring semantic groups of nouns to determine how the logical clustering of nouns affects the overall plausibility of a degree reading. In English, physical objects such as books or vehicles clearly do not have an inherent scale, as opposed to words like “idiot”, “failure” and “effort”.

To address our hypotheses, we construct a logistic regression statistical model from the data we extract from the LASSY corpus.

1.3 Paper Outline

This paper is organized as follows: In Section 2, we discuss the theoretical framework for our analysis, including additional introductory background on the topic of gradability. In Section 3, we discuss the procedure by which we gathered data, and in Section 4 we describe our statistical analysis. In Section 5, we discuss the results of our analysis and give our concluding remarks in Section 6. Section 7 contains a discussion of challenges within our analysis, as well as ideas for future research. Several appendices provide tables containing detailed results of our analysis.

2. Theoretical Framework

In his 2009 paper Morzycki elaborates on the phenomenon of size adjectives modifying the degree of a noun rather than its size, as shown in the example below:

(4) a. A big ship.
b. A big idiot.

The adjective-noun pair in example (4-a) has a normal size reading; the adjective "big" modifies the size of the noun, thus stating that the physical size of the ship is large. (4-b) however, is ambiguous; the size adjective may either modify the physical size of the idiot in question or the degree of his idiocy. The latter seems to be the preferred reading by native English speakers. As shown, only inherently gradable nouns can receive a degree reading.

Morzycki shows that the degree reading is in fact a distinct reading and not just a metaphorical interpretation of the size reading with examples such as the following, which serve as a test for the inherent gradability of nouns:

(5) a. #The ship isn't big, but it's a big ship.
b. The idiot isn't big, but he's a big idiot.

Due to the non-gradable noun "ship" (5-a) is contradictory, while in (5-b) there is no sense of contradiction due to the inherent gradable scale of the noun "idiot".

Morzycki (2009) also contrasts degree readings against what he calls abstract size readings, which are "[abstract] size readings that make reference to size along a possibly abstract dimension – one that may correlate with some intuitive sense of extremeness or severity". Some examples of abstract size readings are listed in (6):

(6) a. A huge thunderstorm.
b. A big concern.
c. A giant disaster.

According to Morzycki these readings are metaphorical in some sense. He pairs them with normal size readings rather than with degree readings and shows that these abstract size readings fail the test shown in (5):

(7) # The thunderstorm wasn't huge, but it was a huge thunderstorm.

In this paper, however, we pair abstract size readings with degree readings, due to the "wh-exclamative" test for inherent gradability, as demonstrated in (8). If a noun is inherently gradable, the wh-exclamative as a whole depicts the object that the noun refers to as one with a high degree of nounhood (for example, (8-a)). If the noun is not gradable, there seems to be an implicit gradable adjective that needs to be inferred from the context (as in (8-b)). This implicit adjective denotes the aspect in which the noun is distinct (Terunuma 1997).

(8) a. What an idiot!
b. What a carpenter!
c. What a salary!

Examples (8-a) and (8-b) above show the difference between degree and quality. The idiot in example (8-a) has a high degree of idiocy; he is likely to be more of an idiot than other idiots. It is unclear, however, in which aspect the carpenter in example (8-b) surpasses other carpenters. This is a pragmatic issue; because the noun is not inherently gradable, context is needed to determine which quality the carpenter possesses that distinguishes him from other carpenters.

In contrast, the salary referenced in (8-c) is not a more prototypical example of a salary, but rather, the exclamative infers that the inherent scale of the noun 'salary' (denoted in amounts of money) is of a noteworthy degree. For this reason, we treat adjective-noun pairs with an abstract size reading in the same manner as pairs with a degree reading.

3. Experimental Framework

This section describes the corpora used in this analysis, as well as the steps taken to format and preprocess the data.

3.1 Gathering Data

To prepare the data for our analysis, we extract phrases with possible gradable readings from LASSY, including syntactic and semantic context; manually annotate each example with a gradable or non-gradable reading; automatically group each noun into semantic groups; and manually revise the automatic noun groups.

Our initial idea was to extract all the adjective-noun pairs containing a size adjective that appear in the LASSY corpus, group the nouns of those pairs into semantic groups, and create a probability distribution for degree readings within these noun groups. Since each example needs to be manually annotated, we restrict our experiment to the adjective-noun pairs containing either "gigantisch", "reusachtig" or "kolossaal" as its size adjective, resulting in 5,874 pairs – a data set that remains large enough to use for statistical analyses.

3.1.1 LASSY Treebank

The data used in our analysis comes from a Dutch corpus called "Large Scale Syntactic Annotation of written Dutch" (LASSY). LASSY was developed as an extension of the Dutch Language Corpus Initiative (D-Coi) under the Flemish-Dutch STEVIN programme by a consortium consisting of the University of Groningen and the Katholieke Universiteit Leuven. The large corpus consists of the CLEF Question Answering Corpus (2005) and the Eindhoven, Mediargus, Senseval, Sonar (release 1), and Twente News (2005) corpora and contains 500 million words.

The LASSY corpus is syntactically annotated according to the Alpino Treebank specification. This format allows the use of XPath and XQuery for linguistically interesting queries on the corpus. A graphical interface, called *dtview* (Data Treebank View) was also built on top of XPath to allow researchers to visualize the dependency structures of sentences within the LASSY corpus and to perform sophisticated linguistic queries.

Extracting data. The following outlines the steps we follow to extract data from LASSY for our analysis.

First, we look for all adjective-noun pairs in the corpus, where the adjective is "gigantisch", "reusachtig", or "kolossaal". We then capture the entire NP constituent by retrieving the parent node of the adjective-noun pair. Once the NP is captured, we extract the adjective's and noun's root forms for later use. If there exists a determiner that immediately precedes the adjective, we capture it to determine if the NP is definite or indefinite. Unspecified determiners are marked as null. If a PP immediately follows the noun, then the example was marked as having an adjunct preposition.

The grammatical roles we are interested in are whether the NP is in subject position, object position, or is part of a predicative complement. These roles are annotated in the LASSY Treebank as *subj*, *obj1*[2], and *predc*, respectively. The grammatical role of the NP is typically annotated directly at the NP's node; however, there are certain cases where the grammatical role is not so clear. If the grammatical role cannot be determined at the NP level, then we perform a cascade search by interrogating the NP's parent node to capture its grammatical role. We recursively search until a valid grammatical role is found, or we reach the root node of the dependency structure – in which case, we mark the grammatical role as *other*.

Figure 1a provides an example of a typical gradable example in LASSY. In this example, *een gigantisch probleem* contains a determiner, but not an adjunct preposition. We record the phrase, along with its determiner, and make a note that there is not an adjunct preposition. We then look at the annotations on the NP to determine its grammatical role. In this scenario, the NP is in object position.

Figure 1b is an example of an NP with an adjunct preposition, but with an undefined grammatical role. We could not identify a valid grammatical role by cascading through the NP's parents, so this example is marked with a grammatical role of *other*.

Figure 1c provides an example where the grammatical role is determined by cascading to the NP's parent node (an AP). The AP node reveals that the NP is in the predicative complement role in the sentence.

Transforming determiner data. As mentioned earlier, the purpose for extracting the determiner type from each NP example was to determine if the NP is definite or indefinite. In Section 1.2, we describe how to determine the type of the article. For our purposes, we stored the definiteness of the NP, as well as the "preceding article type". We divided the NPs into four groups, based on the determiner type: *definite*, *indefinite*, *numeric*, or *null*. NPs with an indefinite, numeric or null determiner are classified as indefinite NPs, NPs with a definite determiner were classified as definite NPs.

3.1.2 Manual Annotations

The next step involves manually annotating every adjective-noun pair that was extracted from the LASSY Treebank for gradability. For every pair it has to be decided whether the size adjective was meant to increase the size of the thing depicted by the noun, or the degree of it. Some pairs are clear cut cases of either a gradable reading or a size reading, but other cases are fairly ambiguous, such as in (9-a).

2. It should be noted that LASSY annotates prepositional complements as well as clausal objects as obj1 . While we are not equating the two roles, we leave them together as we are concerned with finding simple cues that suggest degree readings for noun phrases.

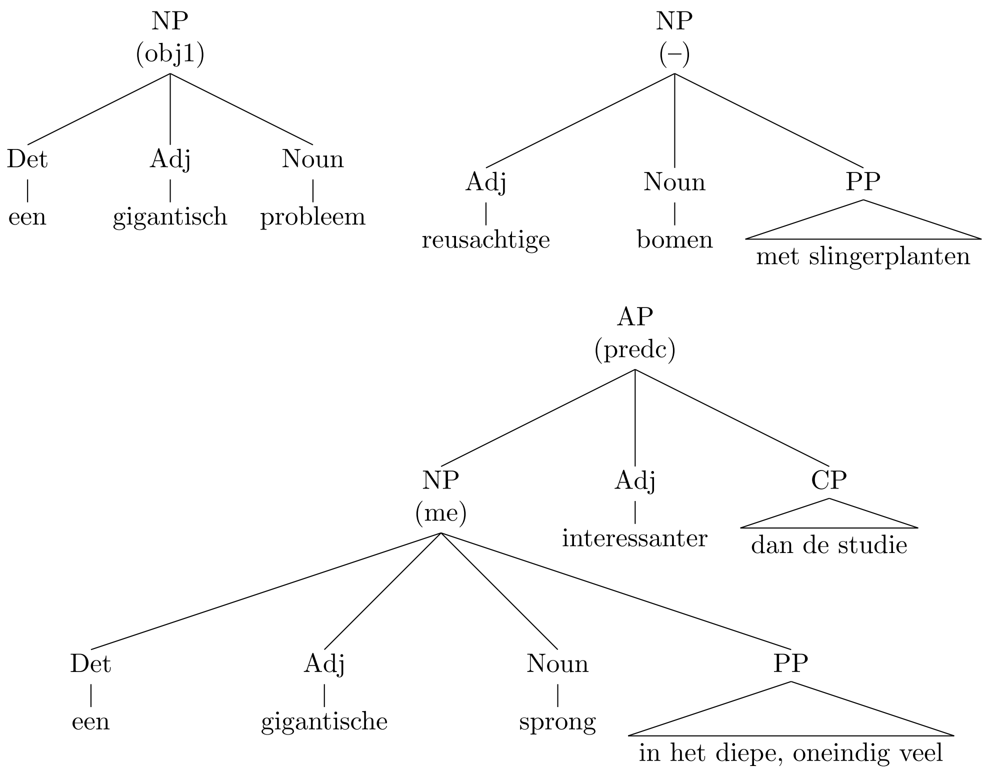
(9) a. Een reusachtige kater
A giant tomcat/hangover

b. Een reusachtige kater na de uitschakeling in de tweede ronde van het toernooi om de Uefa Cup
A giant hangover after the elimination in the second round of the tournament for the Uefa Cup

The Dutch word “kater” has two meanings: a male cat or a hangover. In order to disambiguate nouns, we extract the entire constituent that the pair is part of. As shown in (9-b), it becomes clear that “kater” means hangover and thus contains a gradable scale. In some cases the constituent itself does not provide enough context, in which case we extract the entire sentence or in some cases the entire article that the pair appears in.

Certain pairs remain hard to annotate, regardless of the available context – particularly pairs with nouns that depict abstract things that in some way relate to money like “winstmarge” (profit margin). After much deliberation all such nouns are considered to have an inherent scale resulting in the pairs being annotated as having a degree reading.

(a) *een gigantisch probleem* (b) *reusachtige bomen met slingerplanten*



(c) *een gigantische sprong in het diepe, oneindig veel interessanter dan de studie*

Figure 1: Examples of dependency structures of gradable examples in LASSY.

3.1.3 Automatic Noun Grouping with Cornetto

In order to look at the effect of certain nouns on the plausibility of a gradable predicate, we group the noun pairs by their semantic similarity using Cornetto (Vossen et al. 2008), a lexical hierarchical semantic database for Dutch that covers over 92,000 lemmas as well as various ontologies and relationships, including listings of vertical and horizontal semantic relations. With the use of Cornetto, we build a dictionary of relationships between each extracted noun. From this dictionary, we form noun groups by clustering related nouns. Since many nouns have hyperonymic relations with "iets" (something) within a hierarchical distance as little as 3 levels, we only cluster nouns with direct semantic relationships based on synonymy, hyper-/hyponomy, or meronymy.

Another computational challenge in automatically grouping nouns involves the sheer number of nouns in our LASSY examples. We have a total of 2,710 distinct nouns, which would yield over 7 million noun pairs to cluster. As a result, we prune nouns that occur less than 4 times, which yields 297 distinct nouns. The pruned nouns are grouped into a miscellaneous noun group. To assign nouns to semantic groups, we adopt a "greedy approach", based on adjective-noun pair frequencies, defined in Algorithm 1.

Algorithm 1 Greedy approach to forming automatic noun groups with Cornetto

```
sort nouns by the frequency it appears with an adjective in the LASSY examples (descending
order)
for all nouns with frequency > 4 do
   if noun is not in a noun group then
      create a noun group with the noun as its head
   end if
   for all noun senses in noun do
      find all noun senses with a direct semantic relation in Cornetto
      map all noun senses to the original noun sense
   end for
end for
```

The head of each noun group was the most frequently occurring noun in the group. After completing Algorithm 1, we pruned several infrequent noun groups to avoid data sparseness.

3.1.4 Manual Refinement of Cornetto Groups

The decision to only allow direct semantic relationships combined with the greedy approach used to create groups leads to a large number of semantic groups (some of which seemed quite odd). In order to further restrict the number of noun groups generated from Cornetto and to increase their quality, we manually merge groups and reassigned the noun heads. For example, noun groups of "auto" (car), "boot" (boat), "vliegtuig" (airplane) are merged into a general semantic group of "transportmiddel" (mode of transport).[3]

3. Certainly, it would be more accurate to create the noun groups manually, but to do so would be time-prohibitive. It is more interesting in terms of time savings to augment an automatic approach.

3.1.5 Data Summary

In total, we extract 5,874 phrases from LASSY where a noun was preceded by an adjective of “gigantisch”, “reusachtig”, “kolossaal”. Of the 5,874 examples, 2,161 are identified as having a gradable reading, while the remaining 3,713 have a normal size reading. Table 1 provides breakdowns of gradable and non-gradable readings by adjective choice, grammatical role, and adjunct preposition, respectively.

Adjective	Non-Gradable	Gradable	Total
gigantisch	1482	1492	2974
kolossaal	606	242	848
reusachtig	1625	427	2052
Total:	3713	2161	5874

(a) Gradable Readings by Adjective.

Role	Non-Gradable	Gradable	Total
obj1	2318	1373	3691
other	277	141	418
predc	225	275	500
su	893	372	1265
Total:	3713	2161	5874

(b) Gradable readings by Grammatical Role.

Definite?	Non-Gradable	Gradable	Total
No	2449	1505	3954
Yes	1264	656	1920
Sum	3713	2161	5874

(c) Gradable readings by Definiteness.

Adjunct PP?	Non-Gradable	Gradable	Total
No	2723	1648	4371
Yes	990	513	1503
Total:	3713	2161	5874

(d) Gradable readings by Adjunct Preposition.

Table 1: Gradable reading assignments in the LASSY corpus.

4. Logistic Regression Analysis

We constructed several logistic models to test our research hypotheses and to explore our data. The logistic regression analyses were carried out with the glm function of the “stats” package in R version 2.11.1 in the Windows 7 operating environment.

4.1 Model 1

Our first model was constructed to predict gradable readings based on adjective, grammatical role, manually corrected noun group, definiteness, and the presence of an adjunct PP (true or false). In a logistic regression analysis, the null hypothesis states that none of the factors included in the model are significant.

We first demonstrate the significance of the overall model by comparing it to the intercept-only model, which contains no predictors and simply categorizes all examples as non-gradable. We use a likelihood ratio test for overall model significance. The null deviance of the intercept-only model is 7,728 with 5,873 degrees of freedom and our model has a residual deviance of 4,456 with 5,817 degrees of freedom. Baayen (2008) explains that the difference between the null deviance and the residual deviance approximately follows a chi-square distribution with 56 degrees of freedom. Under the chi-square distribution, we reject the null hypothesis with $p < 0.05$.

Factor	Design Variables
Adjective	2
Grammatical Role	3
Definiteness	1
Adjunct PP	1
Manually-Corrected Noun Group	49
Automatically-Constructed Noun Group	74

Table 2: Number of design variables created for each factor.

We next use an ANOVA analysis to summarize the model and we list the summary in Table 3. The choice of adjective, grammatical role, and noun group, are shown to be highly significant. We shall explore the effects of these predictors later. On the other hand, the definiteness and the presence of an adjunct preposition after a NP have no significant effect on the plausibility of a gradable reading, as the p-values of these columns do not fall within the $p < 0.05$ confidence threshold.

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			5873	7728.12	
Adj	2	492.33	5871	7235.79	0.0000
CorrectedNounGroup	49	2753.41	5822	4482.38	0.0000
GrammaticalRole	3	25.53	5819	4456.85	0.0000
Definite	1	0.45	5818	4456.39	0.5007
FollowingPP	1	0.40	5817	4456.00	0.5285

Table 3: ANOVA analysis of significance for the predictors in Model 1.

4.2 Model 2

We were particularly surprised that, according to Model 1, the definiteness of a NP does not affect the plausibility of a gradable reading. Thus, we decided to explore the matter in more detail by partitioning the predictor further. Recall that we defined four categories for articles that precede the adjective in our example NPs: definite, indefinite, number, and null. In Model 2, we replace the "Definite" predictor with a new predictor, called the "preceding determiner type." This predictor categorizes the determiner into the four aforementioned categories.

Table 9 in Appendix B shows a summary of the statistics in the logistic regression analysis. The fourth column lists the Wald test statistics of each of the design variables constructed from the predictors in our model. The Wald z-score is calculated as the value of the design variable's coefficient, divided by its standard error. As stated in Hosmer and Lemeshow (2000), these Wald statistics follow a standard normal distribution under the hypothesis that a given individual coefficient is zero. Using a level of significance of 0.05, we conclude that any of the design variables that have an associated p-value less than 0.05 (shown in the fifth column) are statistically significant predictors. We again assess the overall utility of our predictors, we use an ANOVA test in Table 4.

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			5873	7728.12	
Adj	2	492.33	5871	7235.79	0.0000
CorrectedNounGroup	49	2753.41	5822	4482.38	0.0000
GrammaticalRole	3	25.53	5819	4456.85	0.0000
PrecedingArticleType	3	27.86	5816	4428.98	0.0000

Table 4: ANOVA analysis of significance for the predictors in Model 2.

As shown in Table 4, the type of determiner is shown to also be highly significant. So why is there a discrepancy between the definiteness of the NP, versus the type of determiner attributed to the NP? Looking at the coefficients of the design variables in Appendix B, we notice that the numeric determiners are the only determiners that are statistically significant ($p < 0.05$).

According to the model, the odds of an NP receiving a gradable reading is affected less by the adjectives "reusachtig" and "kolossaal", as opposed to "gigantisch". All other factors held constant, the odds of an NP with the adjective "gigantisch" are 1.893 times more likely to yield a gradable reading than "kolossaal" ($1/e^{-0.638}$) and 2.555 times than "reusachtig" ($1/e^{-0.938}$).

Our logits for the grammatical role predictor analyze the effect of a specific grammatical role against the default bucket of "other". According to the p-values of the design variables, only the predicative complement position has a significant difference from the "other" category: it is 1.638 times more likely to yield a gradable reading than a NP in the default category ($e^{0.493}$). If a NP has a numeric preceding determiner, it is 0.141 times as likely to be interpreted as a degree reading than other definite determiners ($e^{-1.958}$). The remaining determiner types do not have a significant effect on the plausibility of a gradable reading.

Many of the noun groups used in Model 2 do not have a significant effect ($p < 0.05$); however, the following noun groups have a positive effect in relation to the "other" category: BEDRAG, HOEVEELHEID, KARWEI, KLAP, KLUS, OPERATIE, OPGAAF, STIJGING, TOENAME, VERSCHIL, and WINST. APPARAAT, DIER, DIMENSIE, DING, GEBOUW, LAND, LICHAAM, and PLANT had a negative effect.

We also tested the accuracy of Model 2 by comparing the results of the predictive model to the actual gradability assigned to each example. We assume that any prediction with

$p \geq 0.5$ is classified as a gradable reading, while any p below this threshold is considered to be non-gradable. Table 5 lists the overall classification accuracy of Model 2. A baseline, intercept-only model assuming that all coefficients are 0 has a classification accuracy of 63.21%. Our model has an overall accuracy of 80.23%. Appendix B contains the 95% confidence intervals for the coefficients.

	Actual = 0	Actual = 1	Overall
Predicted = 0	3497	945	4442
Predicted = 1	216	1216	1432
Percent Correct	94.18%	56.27%	80.23%

Table 5: Classification accuracy of Model 2. The overall classification accuracy of Model 2 is 80.23%, as opposed to the null model, which has an accuracy of 63.21%.

Table 6 provides 10 examples from the LASSY corpus, with the estimated likelihoods of each example having a gradable reading. While 9 out of 10 examples classified the gradable reading correctly, the model had difficulties with the phrase "een reusachtig waterballet".

Example #	Context	Predicted p	Actual
2182	een gigantisch project	0.94	1
5404	gigantische vruchten	0.03	0
226	het gigantische openluchttheater van zijn geboortestad St.Louis	0.48	0
3465	een reusachtig verlies	1.00	1
2613	een reusachtige herdenkingskoepel met galerijen	0.26	0
3840	de gigantische post landbouwuitgaven 50 procent van het geheel	0.96	1
5656	kolossale problemen	0.99	1
1913	een reusachtig waterballet	0.26	1
2062	een gigantische wand van gekreukt zilverpapier	0.00	0
782	een gigantisch probleem	1.00	1

Table 6: A random sample of 10 examples in the corpus, with predictions based on Logistic Regression Model 2. Any prediction with $p \geq 0.5$ is considered to be a gradable reading.

5. Results

In Section 4, we provided the results of the statistical analysis for Models 1 and 2, which can be used to address our research hypotheses.

5.1 Hypothesis Testing

Grammatical role. Our first hypothesis was that a NP (containing an adjective-noun pair) in the grammatical role of object or predicative position is more likely to allow a degree reading than a NP in the subject position of a sentence. However, the results of our second regression analysis (Model 2) showed that only the predicative position had a significant effect on the probability of a degree reading ($p > 0.05$). Neither the subject nor the object position had a significant effect. The actual values can be found in Appendix B.

These results have led us to reject our hypothesis that the object and predicative position are more prone than the subject position to receiving a degree reading. As it turns out, the predicative position is more likely to receive a degree reading than the subject and the object position; being in subject or object position has no (significant) effect on the probability of the adjective noun pair to have a degree reading.

Preceding article type. Another hypothesis was that indefinite NPs (containing an adjective-noun pair) are more prone to having a degree reading than definite NPs. However, as presented in Section 4.1, the results of our first logistic regression analysis (Model 1) showed that our variable of definiteness was not even close to being significant ($p = 0.4829$).

This led us to distinguish between four types of determiners; definite, indefinite, number and null (instead of just two types: definite and indefinite) and to do another logistic regression analysis (Model 2). The second analysis revealed that only the numeric determiners have a significant (negative) effect on the probability of an adjective-noun pair to have a degree reading. Definite, indefinite and null determiners do not have a significant effect (the actual values can be found in Appendix B).

So, the results of our analyses reject the hypothesis that indefinite NPs are more prone to having a degree reading. Instead they show that definite, indefinite and null determiners all have no significant effect on the probability to have a degree reading and only numeric determiners have a (negative) effect on this probability.

Adjunct prepositions. We hypothesized that the presence of an adjunct preposition has no effect on the plausibility of an adjective-noun pair receiving a degree reading. The results of our first logistic regression analysis (Model 1), discussed in Section 4.1, show that this variable is far from significant ($p = 0.6113$). Therefore we can accept the hypothesis that an adjunct preposition has no effect on the plausibility of a degree reading.

Noun groups. As shown in Section 4.2 a number of noun groups proved to have a significant effect on the probability of a degree reading, some groups increasing the probability and others decreasing it. Examination of these groups showed that these results are intuitively sound. This section contains a discussion of some of the significant noun groups.

The noun group BEDRAG (sum of money) is made up out of examples that involve money and as was mentioned in 3.1.2, we consider these nouns to have an inherent scale. Below are some examples of pairs that appear in this group:

(10) a. Een reusachtige schuld
   A giant debt
b. Het gigantische prijskaartje
   The gigantic price tag

Clearly the "schuld" in (10-a) cannot be giant in a physical sense; the adjective-noun pair targets a large sum of money that is owed. Likewise, in (10-b) it is not an actual gigantic "prijskaartje", but rather the price on the price tag that is very high (or expensive).

Due to the greedy approach taken in the automatic creation of noun groups (as discussed in Section 3.1.3) the name of the group WINST is somewhat misleading. "Winst" translates to either *profit* or *victory*, but the group is filled with words that depict some form of success. The following examples will give an idea of the content of this noun group:

(11) a. De reusachtige hit
The giant hit(-record)
b. Een kolossale overwinning
A colossal victory

Due to the abstract nature of the nouns in (11) they do not receive a size reading. They are all cases of an abstract-size reading and, as discussed in Section 2, these are considered to have an inherent scale and are thus marked as gradable. Both BEDRAG and WINST had a positive effect on the probability of a degree reading.

A noun group that had a negative effect on the probability of a degree reading was GEBOUW (building). This group is filled what nouns like "kerk" (church), "ziekenhuis" (hospital), "paleis" (palace), etc.

It will not come as a surprise that the adjective-noun pairs in this group are, by and large, clear cut cases of size readings. So the statistically determined negative effect that being a member of this group has on the probability of receiving a degree reading, is intuitively correct, as are the positive effects that the groups BEDRAG and WINST have on this probability.

6. Conclusion

We have presented a method of verifying and extending Morzycki (2009)'s ideas of syntax affecting the plausibility of size adjectives modifying the degree rather than the physical size of a noun through a corpus-based study in the Dutch language. Using the LASSY Treebank and Cornetto thesaurus tool for Dutch, we extracted adjective-noun pairs from dependency structures containing either "gigantisch", "kolossaal" or "reusachtig" and tested the effects of grammatical roles, the definiteness of a determiner, the semantic noun groups, and the presence of an adjunct preposition on the probability of an adjective-noun pair having a degree reading.

By modeling the data with logistic regression, we verified that the grammatical role of predicative complement would be more likely to allow a degree reading than the subject or object positions. We additionally discovered that neither definite, indefinite nor null determiners had a significant effect; only numeric determiners had a significant (negative) effect on the probability of a degree reading.

We also attempted to find out how the logical clustering of nouns would affect the overall plausibility of a degree reading. The results of our analysis revealed that some of the groups that we created indeed had a significant effect. Groups with a positive effect as well as groups with a negative effect on the probability of a degree reading were found.

7. Suggestions for future research

Adjectives. The scope of our research experiment was limited by the lack of annotated data on degree readings in dependency structures. While LASSY contains many other interesting size adjectives, such as "groot" (big) or "enorm" (enormous), over 100,000 instances would need to be annotated. One way to annotate additional adjectives would be to use crowd-sourcing related services, like Amazon's Mechanical Turk, to find annotators to manually and cost-effectively annotate the examples.

There are several approaches to extracting semantic relations from Cornetto. Our "greedy approach" managed to define some interesting noun groups, but the limitations we encountered with semantic distance and our preference for grouping nouns by the most frequent noun in the corpus yielded many groups that needed to be hand-modified. Further work should be performed to determine more intelligent methods to extracting noun groups from Cornetto and identifying the correct word sense of a noun in a given example. For example, Resnik (1995) discusses a method for disambiguating noun groupings with respect to WordNet senses.

Abstract size nouns. Most of the noun groups in our experiment that were found to have a significant positive effect on the probability of a degree reading consist of nouns that do not have a physical size. In our study, we focused on extracting noun phrases with one of three size adjectives, but assumed no constraints on the selection of nouns. In order to study nouns with inherent degree readings better, it is advisable to focus on nouns that are ambiguous with respect to degree readings and physical size readings such as the nouns outlined in Morzycki (2009).

Prediction. We have shown that our logistic regression model has predictive power. With more annotated data, the prediction of degree readings can be useful in sentiment analysis and information extraction scenarios.

Acknowledgments

This work was completed with the support of Jack Hoeksema, John Nerbonne, Gertjan van Noord, and Gosse Bouma from the University of Groningen. The first author received a grant under the Erasmus Mundus Language & Communication Technologies programme.

References

Baayen, R. Harald (2008), *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R.*, Cambridge University Press.

Hosmer, David W. and Stanley Lemeshow (2000), *Applied Logistic Regression (Wiley Series in Probability and Statistics)*, 2 ed., Wiley-Interscience Publication.

Morzycki, Marcin (2009), Degree modification of gradable nouns: size adjectives and adnominal degree morphemes, *Natural Language Semantics* **17** (2), pp. 175–203. http://dx.doi.org/10.1007/s11050-009-9045-7.

Resnik, Philip (1995), Disambiguating Noun Groupings with Respect to WordNet Senses, *Proceedings of the Third workshop on Very Large Corpora*, pp. 54–68. http://acl.ldc.upenn.edu/W/W95/W95-0105.pdf.

Terunuma, Akiko (1997), A note on preposition pied-piping and preposition stranding in wh-exclamatives, *Linguistic research : working papers in English linguistics* **15**, pp. 53–64, Tokyo University English Linguistics Association, University of Tokyo. http://ci.nii.ac.jp/naid/120000873542/en/.

Vossen, Piek, Isa Maks, Roxane Segers, and Hennie van der Vliet (2008), Integrating Lexical Units, Synsets and Ontology in the Cornetto Database, *in* Nicoletta Calzolari (Conference Chair), Bente Maegaard Joseph Mariani Jan Odjik Stelios Piperidis Daniel Tapias, Khalid Choukri, editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, European Language Resources Association (ELRA), Marrakech, Morocco. http://www.lrec-conf.org/proceedings/lrec2008/.

Appendix A. Noun Groups in our analysis

Tables 7 and 8 list the noun groups that were constructed for the statistical analysis. Note that the noun groups are listed with a noun sense. To see how each noun sense is used, please refer to the Cornetto database.

Table 7 lists the pruned noun groups that were automatically constructed with Cornetto. Section 3.1.3 describes the automatic grouping process. Table 8 lists the pruned noun groups that were constructed by manually editing the noun groups generated with the help of Cornetto.

Appendix B. Logistic Regression

Table 9 lists the coefficients obtained in the logisitic regression analysis for Model 2. The coefficients with an associated probability less than 0.05 are considered to be statistically significant in our analysis to maintain a 95% confidence interval. Due to space limitations, we do not include the coefficients for models 1 and 3.

Table 9: Logistic Regression Analysis for Model 2: Using manually corrected noun groups and 4 distinct determiner types.

	β	S.E.	z value	Pr(>|z|)
(Intercept)	-0.1246	0.1527	-0.82	0.4145
Adj:kolossaal	-0.6382	0.1095	-5.83	0.0000
Adj:reusachtig	-0.9381	0.0844	-11.12	0.0000
afmeting:1	-1.2104	0.7729	-1.57	0.1174
apparaat:1	-3.2310	1.0142	-3.19	0.0014
bak:1	-18.0879	1185.5103	-0.02	0.9878
bedrag:1	3.3284	0.2202	15.12	0.0000
bedrijf:1	-18.1380	1463.0176	-0.01	0.9901
beeld:1	-17.7140	1609.4306	-0.01	0.9912
beeldscherm:1	-17.9051	799.4170	-0.02	0.9821
billboard:1	-17.8237	1548.5533	-0.01	0.9908
chaos:1	0.8835	0.5104	1.73	0.0835
complex:1	-18.0969	1389.1865	-0.01	0.9896

Table 9 – Continued

	β	S.E.	z value	Pr(>|z|)
dier:1	-4.5101	1.0043	-4.49	0.0000
dimensie:1	-3.1202	0.7211	-4.33	0.0000
ding:1	-2.3837	0.5182	-4.60	0.0000
doek:1	-17.7001	1443.3955	-0.01	0.9902
ervaring:2	18.9154	1704.0481	0.01	0.9911
foto:1	-17.6368	1531.9178	-0.01	0.9908
gat:1	-18.3390	1309.5746	-0.01	0.9888
gebouw:1	-5.5385	1.0024	-5.52	0.0000
gevolg:1	18.8309	1565.9452	0.01	0.9904
hoeveelheid:1	1.4568	0.1351	10.78	0.0000
karwei:1	2.7691	0.5253	5.27	0.0000
klap:1	1.4931	0.3619	4.13	0.0000
klus:1	1.7789	0.6058	2.94	0.0033
kunst:1	-17.7716	521.3477	-0.03	0.9728
land:1	-2.9096	1.0256	-2.84	0.0046
lichaam:1	-3.0342	0.5872	-5.17	0.0000
lichaam:3	-1.9903	1.0392	-1.92	0.0555
machine:1	-17.8528	903.0852	-0.02	0.9842
manier:1	0.8619	0.4266	2.02	0.0433
man:3	-0.2419	0.4904	-0.49	0.6218
netwerk:1	-18.2349	1559.5637	-0.01	0.9907
onderneming:1	1.2278	0.5810	2.11	0.0346
operatie:1	1.5524	0.4623	3.36	0.0008
opgaaf:1	4.9902	1.0071	4.96	0.0000
partij:4	-1.0056	0.5231	-1.92	0.0546
persoon:1	-0.2994	0.2615	-1.15	0.2522
plant:1	-3.3997	1.0107	-3.36	0.0008
publicatie:2	-1.2003	0.6245	-1.92	0.0546
ruimte:2	-18.1045	632.5243	-0.03	0.9772
stap:1	-0.4753	0.4794	-0.99	0.3215
stijging:1	3.1412	1.0384	3.03	0.0025
stuk:1	-1.2740	0.7752	-1.64	0.1003
toename:1	1.0964	0.4408	2.49	0.0129
transportmiddel:1	-17.8129	590.2721	-0.03	0.9759
verlies:1	19.1263	1051.7269	0.02	0.9855
verlies:3	18.9735	1078.7608	0.02	0.9860
verschil:1	1.4140	0.4635	3.05	0.0023
vorm:1	-0.1524	0.5544	-0.27	0.7834
winst:3	3.8648	0.7233	5.34	0.0000
GrammaticalRole:obj1	0.0341	0.1421	0.24	0.8106
GrammaticalRole:predc	0.4932	0.1849	2.67	0.0077
GrammaticalRole:su	-0.2531	0.1583	-1.60	0.1098
PrecedingArticle:indefinite	-0.0123	0.0860	-0.14	0.8862
PrecedingArticle:none	-0.0299	0.1076	-0.28	0.7813
PrecedingArticle:number	-1.9580	0.4691	-4.17	0.0000

Noun Group	Members
hoeveelheid:1	265
bedrag:1	210
gebouw:1	157
probleem:1	97
lichaam:1	90
man:3	82
afmeting:1	76
ding:1	72
stuk:1	72
klus:1	68
succes:1	63
apparaat:1	53
aantal:1	52
verzameling:1	42
vorm:1	41
boom:1	40
verlies:1	37
beeld:3	36
lichaam:2	36
ruimte:2	36
land:1	35
beeld:1	35
operatie:1	33
beeld:2	30
schip:1	30

Noun Group	Members
lichaam:3	27
boek:1	24
investering:1	24
verschil:1	24
aanbod:1	23
fout:2	23
tent:3	23
complex:1	21
hit:1	21
insect:1	21
portret:1	21
salaris:1	21
scherm:3	20
stap:1	20
wolk:1	20
dier:1	19
foto:1	19
doek:1	18
project:1	18
tuin:1	18
hal:2	17
tekort:3	17
markt:1	16
netwerk:1	16
klap:1	15

Noun Group	Members
terrein:1	15
verlies:2	14
bouw_put:1	13
consequentie:1	13
doek:3	13
explosie:1	13
hand:1	13
aanbod:2	13
belasting_verlaging:1	12
risico:1	12
groei:1	11
huis:1	11
inspanning:1	11
schaal:1	11
video_scherm:1	11
belang:1	10
feest:1	10
kathedraal:1	10
kop:1	10
prijs:1	10
puinhoop:1	10
rij:1	10
schuldenlast:1	10
vraag:4	10
OTHER	3304
Total:	5874

Table 7: Greedily assigned noun groups, using Cornetto.

Noun Group	Members
bedrag:1	438
gebouw:1	435
hoeveelheid:1	325
dier:1	193
kunst:1	148
lichaam:1	138
transportmiddel:1	116
opgaaf:1	107
ruimte:2	101
ding:1	82
persoon:1	82
winst:3	79
dimensie:1	75
plant:1	68
beeldscherm:1	63
karwei:1	58
apparaat:1	54
machine:1	50
klap:1	46
verlies:1	37
verlies:3	35
operatie:1	32
bak:1	29
land:1	26
manier:1	25
publicatie:2	25
gat:1	24
verschil:1	24
toename:1	23
man:3	22
complex:1	21
partij:4	20
stap:1	20
bedrijf:1	19
doek:1	19
klus:1	19
chaos:1	18
billboard:1	17
foto:1	17
gevolg:1	17
netwerk:1	17
stijging:1	17
beeld:1	16
lichaam:3	16
vorm:1	16
afmeting:1	14
ervaring:2	14
onderneming:1	14
stuk:1	14
OTHER	2589
Total:	5874

Table 8: Manually corrected noun groups.

Table 10: 95% confidence level estimations for parameters in Model 2.

			Logit		Odds Ratio	
	β	S.E.	2.5 %	97.5 %	2.5 %	97.5 %
(Intercept)	-0.125	0.153	-0.426	0.174	0.653	1.190
Adj:kolossaal	-0.638	0.110	-0.854	-0.425	0.426	0.654
Adj:reusachtig	-0.938	0.084	-1.104	-0.773	0.332	0.461
afmeting:1	-1.210	0.773	-3.082	0.112	0.046	1.119
apparaat:1	-3.231	1.014	-6.109	-1.701	0.002	0.182
bedrag:1	3.328	0.220	2.920	3.787	18.541	44.121
chaos:1	0.884	0.510	-0.083	1.957	0.921	7.078
dier:1	-4.510	1.004	-7.378	-3.014	0.001	0.049
dimensie:1	-3.120	0.721	-4.931	-1.952	0.007	0.142
ding:1	-2.384	0.518	-3.579	-1.493	0.028	0.225
gebouw:1	-5.538	1.002	-8.405	-4.049	0.000	0.017
hoeveelheid:1	1.457	0.135	1.196	1.726	3.306	5.617
karwei:1	2.769	0.525	1.860	3.974	6.423	53.204
klap:1	1.493	0.362	0.818	2.250	2.265	9.485
klus:1	1.779	0.606	0.683	3.113	1.980	22.492
kunst:1	-17.777	521.974	-229.019	-125.504	0.000	0.000
land:1	-2.910	1.026	-5.798	-1.341	0.003	0.262
lichaam:1	-3.034	0.587	-4.439	-2.055	0.012	0.128
man:3	-0.242	0.490	-1.286	0.672	0.276	1.958
manier:1	0.862	0.427	0.046	1.740	1.047	5.696
onderneming:1	1.228	0.581	0.117	2.446	1.124	11.541
operatie:1	1.552	0.462	0.711	2.555	2.036	12.876
partij:4	-1.006	0.523	-2.139	-0.043	0.118	0.958
persoon:1	-0.299	0.261	-0.831	0.198	0.435	1.219
plant:1	-3.400	1.011	-6.274	-1.882	0.002	0.152
publicatie:2	-1.200	0.625	-2.654	-0.119	0.070	0.888
ruimte:2	-18.104	632.508	-260.972	2.634	0.000	13.928
stap:1	-0.475	0.479	-1.473	0.439	0.229	1.551
stijging:1	3.141	1.038	1.531	6.042	4.624	420.707
stuk:1	-1.274	0.775	-3.148	0.056	0.043	1.058
toename:1	1.096	0.441	0.246	1.997	1.278	7.364
transportmiddel:1	-17.813	590.253	-248.944	0.721	0.000	2.057
verschil:1	1.414	0.463	0.544	2.389	1.724	10.904
vorm:1	-0.152	0.554	-1.332	0.890	0.264	2.435
winst:3	3.865	0.723	2.692	5.679	14.759	292.596
GrammaticalRole:obj1	0.034	0.142	-0.243	0.314	0.784	1.370
GrammaticalRole:predc	0.493	0.185	0.132	0.857	1.141	2.356
GrammaticalRole:su	-0.253	0.158	-0.562	0.058	0.570	1.060
PrecedingArticle:indefinite	-0.012	0.086	-0.181	0.156	0.835	1.169
PrecedingArticle:none	-0.030	0.108	-0.241	0.181	0.786	1.198
PrecedingArticle:number	-1.958	0.469	-2.995	-1.126	0.050	0.324