Abstract

In Dutch, subjects are preferably placed in sentence-initial position. Bare plural subjects, however, are not uncommon in a postverbal position. We will present the results of two studies in which participants had to construct sentences by dragging a subject NP (either bare or definite) or a PP into the pre- or postverbal position. To explain the pattern of word order variation that we found, we analysed the data with an Optimality Theory framework using weighted constraints. We argue that the word order variation with bare plural subjects results from a conflict between constraints that relate to the *Subject First* and the *Topic First* preference. Whereas definite subjects are good topics, causing the two constraints to converge, bare plural subjects are inherently bad topics, causing competition between the constraints. If, however, a compelling context is present, this effect of specificity may be overruled.

*Keywords:* language production, word order, Dutch, optimality theory, bare plurals

# Introduction

Language production is a fast and highly dynamic process. An important question is how seemingly unordered and simultaneous thoughts are transformed into linear structures that characterize language (McDonald et al. 1993). In an incremental view on language production (Kempen and Hoenkamp 1987), utterances are articulated as soon as the first information becomes available for production. Speakers do not postpone their speech until the planning of an utterance is completely finished.

However, the information that becomes available first may not always be the most suitable to start an utterance with, since word order is also constrained by language-specific grammatical rules and discourse requirements. In this paper, we aim to explain the interplay between different mechanisms influencing word order in Dutch. These mechanisms involve factors relating to information structure, such as discourse restrictions and prototypical topic properties, as well as preferences relating to grammatical structure. More specifically, we focus on the position of bare and definite plural subjects in Dutch main clauses.

Dutch is a language with a relatively free word order. It is a verb-second language, in which at most one constituent may precede the finite verb in declarative main clauses. As in many other languages, there is a general preference to place the subject in this preverbal position (Van Tiel and Lamers 2007). However, other constituents, such as direct or indirect objects, oblique complements or adjuncts, that are in some way prominent in the sentence may also occupy this position, causing the subject to be postverbal (Bouma 2008), as illustrated in (1).

(1) *In deze kroeg komen veel taalkundigen.*

 in this pub come many linguists

 ‘Many linguists come to this pub.’

The preverbal position is commonly assumed to be a topic position. It follows that constituents that are prototypically good topics are more likely to be placed in the preverbal position. A prototypical topic is generally a highly accessible entity, for instance because of its animacy or definiteness/specificity (Brunetti 2009; Givón 1976). Many studies have shown that entities that are animate or definite are more likely to be placed in a prominent position (e.g. Van Bergen and De Swart, 2010; Branigan et al. 2008; Prat-Sala and Branigan 2000).

In many cases, the topic of a sentence coincides with the subject. A subject is often interpreted as the topic, because both entities often share properties such as definiteness and agentivity, as has frequently been observed (e.g. Brunetti 2009; Chafe 1976; Lambrecht 1994; Reinhart 1982). Since both subjects and topics are preferred in the preverbal position in Dutch, a subject that is a prototypical topic is expected to be placed in that position. However, when a subject lacks prototypical topic properties, the preverbal position is only preferred on the basis of grammatical function. This may give rise to competition between preferences as to where to place the subject. One preference favors placement of the subject in its canonical preverbal position (henceforth subject-first preference, see Lamers and De Hoop 2007); the other favors the placement of a prototypical topic in the preverbal position (henceforth topic–first preference). If the subject is not a prototypical topic, these two preferences are not in accordance. Especially when another constituent than the subject is present that bears prototypical topic properties, a conflict may arise between grammatical function (start the sentence with the subject) and information structure (start the sentence with the topic), which may influence speakers’ word order choices. This may result in increased variation in constituent linearization.

 In this paper, we focus on word order variation in Dutch that arises with bare (indefinite) plural subjects. Bare plural subjects can be considered prototypically bad topics due to their inability to refer to specific entities. Firstly, we compare the placement of these subjects with that of definite plural subjects, which are considered to be good topics. To ensure that word order variation is possible, we investigate sentences that also include a locative PP adjunct. Locative PPs often appear in the preverbal position in Dutch. They are good topics, because they provide a situational setting for the sentence (e.g. Crasborn et al. 2009). Secondly, we investigate how discourse requirements affect subject placement through their influence on information structure, and how this interacts with preferences in word order based on grammatical function. Three main questions will be addressed: (1) Given that bare plurals lack a prototypical topic property, does the conflict between the subject-firstpreference and the topic-firstpreference have a demonstrable effect on speakers’ placement of bare plural subjects? (2) Can the influence of prototypical topic properties be overruled by constraints of the discourse context on the information structure of the sentence? (3) Is it possible to predict the frequency distributions of different linearizations in language production using a model that implements word order preferences as probabilistic constraints on language production?

 To answer the first two questions, we will present data from two sentence production experiments (Sections 2 and 3). The first experiment addresses research question (1). In this experiment, participants were asked to build sentences using the words provided in a prompt, but without any accompanying context. The second experiment addresses research question (2). In this experiment, the sentences that were produced by the participants were answers to a *wh-*focus question. In Section 4, we present a probabilistic Optimality Theory model based on the experimental data. This makes it possible to analyze the effects of grammatical function, topic properties and information structure on word order in Dutch using weighted constraints. This model will be used to answer research question (3). A general discussion is provided in Section 5. Finally, conclusions are drawn in Section 6.

# Experiment 1: a drag and drop sentence production study without context

To answer the question whether a lack of prototypical topic properties in subjects causes a conflict between the subject-first and the topic-firstpreference, which in turn affects the placement of subjects, we performed a drag-and-drop sentence production experiment. In this experiment, participants constructed sentences which contained either a bare plural subject, as in (2), or a definite plural subject, as in (3).

(2) a. *Kinderen spelen in de tuin.*

 children play in the garden

 ‘Children play in the garden.’

 b. *In de tuin spelen kinderen.*

 in the garden play children

 ‘Children are playing in the garden.’

(3) a. *De kinderen spelen in de tuin.*

 the children play in the garden.

 ‘The children are playing in the garden.’

 b. *In de tuin spelen de kinderen.*

 in the garden play the children

 ‘In the garden, the children are playing.’

In (2a), the bare plural subject *kinderen* ‘children’ is in the preverbal position, in accordance with the subject-first preference. However, since a bare noun is not a prototypical topic, while there is a locative PP adjunct available that could serve as the topic, this sentence violates the topic-first preference. Conversely, in (2b), the bare plural subject is in a postverbal position, the topic position being occupied by the locative adjunct. This word order is in accordance with the topic-first preference, but violates the subject-first preference. As for the definite plural subjects in (3), these are inherently good topics, and hence there is no a priori reason for a definite plural subject to be placed outside its canonical position. Thus, we predict that for bare plural subjects both word orders are equally unmarked in a neutral context, and that word order variation will arise when a sentence with a bare plural subject has to be formulated. For definite plural subjects, we predict that only the preverbal position yields an unmarked word order, and that the preverbal position will be preferred, the postverbal position probably needing a contrastive context to be acceptable. Experiment 1 tested these predictions.

## Method

### Participants

Sixteen native speakers of Dutch (9 male; mean age 22.5 years) participated in the experiment.

### Materials

We created 28 stimuli, consisting of an intransitive verb, a plural animate NP and a locative PP. Each stimulus came in two versions, one in which the NP was a bare plural (e.g. *filosofen* ‘philosophers’) and one in which it was a definite plural (e.g. *de filosofen* ‘the philosophers’), forming the two conditions. Care was taken that sentences created with the three constituents were grammatical both when the NP and when the PP occupied the preverbal position, i.e. in the NP-V-PP order (e.g. *(De) filosofen mediteren in het klooster* ‘(The) philosophers are meditating in the monastery’) and in the PP-V-NP order (e.g. *In het klooster mediteren (de) filosofen* ‘In the monastery (the) philosophers are meditating’). In addition, 32 filler items unrelated to the experimental items and 8 response items (to be explained below) were created.

### Procedure

Participants were seated in front of a computer screen in an experimental room. They were instructed to build grammatically correct Dutch sentences using the three constituents appearing on the screen. The experiment started with the text *Let op! Het gaat nu beginnen!* (‘Attention! It starts now!’) appearing on the screen. Next, a fixation cross appeared for 800 ms in the middle of the screen. Then, the verb was presented in the same position, with dotted lines on both sides to indicate the pre- and postverbal slots. After 1500 ms, the NP and the PP were presented, positioned right above each other in the upper part of the screen. The order of the two constituents was varied, to control for an effect of the order of presentation. Participants had to pick up either the NP or the PP with the mouse, and drop it in the desired position on either side of the verb. The other position was automatically filled with the remaining constituent. The next item followed automatically. A participant had 6000 ms to complete a sentence.

 To isolate influences on word order from discourse related factors, the items were presented without any preceding context. In addition, care was taken that subsequent items were not thematically related. To check whether participants were performing the task attentively, some of the filler items were followed by a word in capitals immediately after a sentence was formed (the response items). Participants had to indicate whether the word had occurred in the sentence they just built by clicking with the mouse on the *ja* ‘yes’or *nee* ‘no’ boxes on the screen. The stimuli were presented in white on a black screen, using the software program *Presentation* (Neurobehavioral Systems[[1]](#footnote-1)). A single experiment consisted of 2 blocks, separated by a short pause. Each experiment was introduced with a practice block of 14 items that were similar to the items in the experimental blocks. It took approximately 15 minutes to complete the experiment.

### Design

The experimental items were distributed over two lists, in such a way that from each set one list contained an item with a bare plural subject in a certain order of presentation and the other an item with a definite plural subject in the reverse order of presentation. In addition, the filler and response items were added to each list. The two lists were divided into two blocks, taking care that items from the same set were kept as far apart from each other as possible.

 A univariate repeated measurement ANOVA was performed, with NP type (bare or definite) as within-subjects factor and stimulus list as between-subjects factor. The dependent variable was the proportion of NP-V-PP word orders.

## Results

The results are shown in Table 1.[[2]](#footnote-2) As can be seen, for both types of subject the preverbal position is most frequent. However, for bare plural subjects the postverbal position is more frequent (42.6%) than for definite plural subjects (18.9%). The effect of type of subject (bare plural or definite plural) on word order was significant (*F*11,11 = 9.37; *p* < .05; *F*21,14 = 60.04; *p* < .001).

(Table 1 about here)

## Discussion

The fact that both bare and definite plural subjects were most frequently placed in the preverbal position in Experiment 1 supports the hypothesis that the subject-first preference is an important constraint on the production of Dutch main clauses. More importantly, the finding that the prevalence of the preverbal position was less strong for bare plural subjects than for definite subjects indicates that the subject-first preference receives competition from anotherpreference with bare plural subjects, resulting in word order variation.

 By not providing any discourse context and controlling the topic properties of the PP by using the same PPs in both conditions, the word order differences we found could only be caused by the definiteness/specificity of the subject NP. Therefore, we argue that the word order variation found for bare plural subjects is caused by competition between the two routes available for subject placement in Dutch. Following the information structure route, a bare plural subject is not likely to be selected as a topic and hence a more accessible constituent will be preferred in the preverbal position. In the grammatical function route, a bare plural subject will be preferred in the preverbal position because of its syntactic function. For definite plural subjects, which are better topics, there is no competition since the two routes converge to the preverbal position.

# Experiment 2: a drag and drop sentence production study in a constraining discourse context

Experiment 1 provided evidence for the influence of specificity as a prototypical topic property on the placement of the subject in Dutch ‘out-of-the-blue’ sentences (without discourse context). However, in the presence of a discourse context that imposes a certain information structure, it might well be that it affects the influence of specificity. For example, if one constituent in the sentence conveys old information, this constituent is preferably placed in the preverbal position, since old information tends to be presented before new information (Haiman 1983). Alternatively, people may want to emphasize new information, because it is important information. To this end, new information may be fronted (focus preposing; Ward 1985), causing the given information to be produced postverbally. The question arises how these effects of the discourse interact with the influence of specificity as a topic property as found in Experiment 1 (without discourse context). This question is addressed in Experiment 2 by adding a *wh* focus question before each stimulus.

## Method

### Participants

Participants were the same as those that participated in Experiment 1.

### Materials

The same stimulus items as in Experiment 1 were used, but a *wh* focus question was added preceding each item. The question either addressed the NP (‘who-question’, e.g. *Wie mediteren in het klooster?* ‘Who are meditating in the monastery?’), or the PP (‘where-question’, e.g. *Waar mediteren (de) filosofen?* ‘Where do (the) philosophers meditate?’[[3]](#footnote-3) In addition, 64 filler items and 12 response items were created.

### Procedure

The procedure was very similar to that of Experiment 1. Only differences will be reported. As in Experiment 1, participants had to build sentences out of three given constituents. However, each sentence was now preceded by a *wh* focus question (either a who-question or a where-question). Participants were instructed to carefully read the question. After 1500 ms, the verb with the dotted lines and the NP and the PP appeared at the top of the screen. The sentence to be built had to be an answer to the question. Participants had 5000 ms to complete a sentence, which was 1000 ms shorter than in Experiment 1, because the participants already knew with which elements the sentence had to be built from the question. Each participant saw 1 stimulus list of 4 blocks. As in Experiment 1, participants started with a short practice block. It took approximately 25 minutes to complete the experiment.

### Design

Crossing the factors question type (who-question or where-question) and NP type (bare or definite) resulted in a 2 x 2 design with four conditions. The stimuli were distributed over two lists, such that the different conditions were spread equally over the lists, all verbs, NPs, PPs and question types occurred equally often on each list and each item occurred only once on one of the lists. Thus, from each set four items occurred on a given list, two with a bare NP and two with a definite NP; two with a who-question and two with a where-question. Care was taken that the order of presentation of the NP and the PP in the four items differed and that each order occurred equally often on a given list. The filler and response items were added to each list, which was subsequently divided into four blocks of 44 items, keeping items from the same set as far apart as possible.

A repeated measurement ANOVA was performed, with NP type and question type as within-subjects factors and stimulus list as between-subjects factor. The dependent variable was the proportion of NP-V-PP word orders.

## Results

(Table 2 about here)

The results are presented in Table 2.[[4]](#footnote-4) The same preference as in Experiment 1 to put the NP in first position, irrespective of definiteness, was observed. However, in contrast to Experiment 1, we found no significant effect of NP type (*F*11,12 = 0.40, *p* = .54; *F*21,14 = 0.15, *p* = .70), suggesting that the effect of definiteness on subject position disappears when the discourse requires either the subject or the PP to be the topic. In addition, the effect of question type was only significant over items, and not over participants (*F*11,13 = 0.73, *p* = .41; *F*21,14 = 83.11, *p* < .001), indicating that this effect cannot be generalized across participants.

## Discussion

The results of Experiment 2 no longer showed an effect of NP type (bare or definite) on word order. As opposed to Experiment 1, there was no difference between the occurrences of bare subjects and definite subjects in pre- or postverbal position. This suggests that the inherent property of non-specificity of bare NPs is overruled by discourse requirements in terms of given and new information.

 As discussed above, new information is likely to be placed in a non-prominent, postverbal position, while given information is likely to be placed in the more prominent preverbal position. Alternatively, new information may be emphasized by putting it in the preverbal position. In the experiment, the new information was the crucial information for the question to be answered, and hence may be assumed to be important information for the (hypothetical) addressee. It is this latter pattern that emerges from the word order frequencies in Table 1: In the answer to a *who*-question, the subject is the new information, and it is more likely to be placed in the preverbal position (NP-V-PP word order) than the subject of an answer to a *where*-question, which is always given information. The fact that this difference was only significant over items suggests that there is a lot of variation between participants in where they place old and new information. When we examined the individual responses of the participants more closely, we found that 14 out of 16 tended to place new information in the preverbal position, while 2 participants mostly put old information in the preverbal position. This suggests that there is variation between people in the strategy they choose to solve the linearization problem, either adhering to the iconic ‘old precedes new’ order, or emphasizing the important information by mentioning the new information first (cf. Givón 1988; Gundel 1988).

 Nevertheless, it is clear that the preference to start a sentence with the subject remains strong. In Experiment 2, as in Experiment 1, the majority of the sentences produced were subject-initial, regardless of the discourse context. This suggests that although context may overrule effects of prototypical topicality, it does not completely override the subject-first preference.

# An OT model with weighted constraints

In this section, we present an Optimality Theory model that predicts the frequencies with which the two possible word orders in intransitive main clauses in Dutch with a plural subject and a locative PP adjunct (NP-V-PP or PP-V-NP) are produced. These predictions are based on the word order preferences and their interplay emerging from the two drag-and-drop experiments. Optimality Theory (OT; Prince and Smolensky 1993) is a theoretical framework in which for a given input the optimal output is selected from a set of possible candidates. This is done by evaluating these candidates against a set of universal but violable constraints. These constraints have language-particular rankings and can be in conflict, in which case a violation of a stronger constraint overrides a violation of a lower ranked constraint.

We use a variant of Optimality Theory in which the constraints are assigned weights, representing their relative strengths. In traditional OT, a candidate is ruled out when it violates the highest constraint that is not violated by another candidate. Whether this suboptimal candidate satisfies any lower ranked constraints that might be violated by other candidates is not relevant for its predicted occurrence. Thus, traditional OT employs an all-or-nothing principle, in which only the optimal candidate is predicted to occur (within a certain context). A model with weighted constraints, by contrast, does not rule out suboptimal candidates, but predicts the frequency of occurrence of each candidate based on the relative strengths of the constraints represented by their weights. As such, it can account for statistical variation in the data.

 There are different ways in which constraint weights, and therefore their rankings, can be determined automatically. In Stochastic OT (Boersma and Hayes 2001), the Gradual Learning Algorithm is used, which takes frequency data as input to learn the position of a constraint on a continuous ranking scale. Constraints that are positioned close enough to each other on this scale may occasionally switch places, thus accounting for statistical variation. Other methods for data-driven constraint ranking include Maximum Entropy models (Goldwater and Johnson, 2003), Harmonic Grammar (Smolensky and Legendre 2006) and Linear Optimality Theory (Keller 2006).

 The main difference between Stochastic OT and the other models is that the latter are trained on previously collected (corpus) data instead of being adapted ‘on line’ based on newly processed observations (Jäger 2007). We follow up on these models by implementing a genetic algorithm (Goldberg and Holland, 1988; Pulleyblank and Turkel 1998) for assigning weights to the constraints in our model. In contrast to Stochastic OT, variation does not stem from occasional switches in the ranking order of the constraints. Rather, violations of lower ranked constraints affect a candidate’s frequency of occurrence, even when a violation of a higher ranked constraint would render this candidate suboptimal in standard OT.

## The constraints

We started our OT model by deriving the necessary constraints from the word order preferences that emerged from the drag-and-drop experiments. The subject-first preference (i.e. the preference to start a sentence with the subject) is captured by the markedness constraint in (4), which expresses that subjects should be avoided in a postverbal position.

(4) \*Post/Subj avoid postverbal subjects

The constraint in (5) relates to a constituent’s topic properties and the preference to place constituents that are prototypically good topics, such as definite NPs and spatio-temporal adjuncts, in the prominent, preverbal position.

(5) \*Post/Top avoid postverbal topics

Note that these constraints combine syntactic position with grammatical function and information structure via harmonic alignment of universal scales of prominence (e.g. Aissen 1999, 2003). Subjects are more prominent than non-subjects (Keenan and Comrie 1977; Lamers and De Hoop 2005), topics are more prominent than non-topics (Givón 1976), and the preverbal position is a more prominent position than postverbal positions (Yang and Van Bergen 2007). By harmonically aligning these scales, subjects and topics in preverbal position come out as more harmonic (unmarked) than subjects and topics in postverbal position.

 Additionally, we want to express that what will be the topic not only depends on prototypical topic properties (Experiment 1), but also on properties derived from the context (Experiment 2). To this end, we formulated the constraints given in (6) and (7).

(6) \*Pre/New avoid new information in preverbal position

(7) FirstThingsFirst give the most important information first

The constraint in (6) expresses the givenness hierarchy, in which given information precedes new information. The constraint in (7) captures the tendency to emphasize important (new) information by placing it in the preverbal position.[[5]](#footnote-5)

## Ranking the constraints

After defining the constraints, the correct ranking of these constraints for Dutch has to be determined. To this end, we trained a genetic algorithm on the data from Experiments 1 and 2 to find the optimal weights of the constraints. When implemented in OT, these weights represent the relative importance of the constraints, and will yield a predicted frequency distribution of the two possible word orders. The ability of a genetic algorithm to find an optimal combination of integers in a large search space, makes it suitable for the current task.[[6]](#footnote-6) After training, the weights were evaluated against unseen data in the test phase. The method we used for training and testing was tenfold cross-validation. In tenfold cross-validation, the data are split up in ten parts of equal size. Training is done on nine parts, while the remaining part is left for testing. This procedure is repeated ten times, each time a different part being used for testing. The final prediction accuracy of the model is calculated by taking the mean of the proportions of correctly predicted data from all runs.

 Six scenarios were distinguished, resulting from the two experiments (i.e. Experiment 1 without context, and Experiment 2 in which a *wh*-focus question was presented):

1. No context, bare plural (Tableau 1)
2. No context, definite plural (Tabl. 2)
3. Who-question, bare plural (Tabl. 3)
4. Who-question, definite plural (Tabl. 4)
5. Where-question, bare plural (Tabl. 5)
6. Where-question, definite plural (Tabl. 6)

To incorporate the influence of the presence of a discourse context on the constraints we separated the data in two batches, one for scenarios 1 and 2 and one for scenarios 3 to 6. This allows for the possibility that constraint weights may vary for scenarios without a preceding context and for those that are preceded by a *wh*-question (cf. Keller 2006).

 We defined the range of possible constraint weights as all integers between 0 and 100. In the case of four constraints, the number of possible weight combinations is 100\*100\*100\*100=108. The genetic algorithm searches through these combinations to find the one that is optimal in relation to a given *fitness function*. The definition of the fitness function for our algorithm is as follows. First, we defined for each word order candidate $A\_{i} ϵ ${NP-V-PP,PP-V-NP}[[7]](#footnote-7) which constraints are violated when this candidate is chosen and which are not. This gives us for each candidate a constraint violations vector $(V\_{1,}..V\_{4})$, in which each $V\_{x} ϵ \{0,1\}$, with 1 meaning that the corresponding constraint is violated. Thus, for example, for candidate A1 = NP-V-PP in scenario 5 (‘where-question, bare plural’), we get the violations vector (0,1,0,1) (see Tableau 5).

 For every weight vector generated by the algorithm $(W\_{1}..W\_{m})$, we calculate the total violation of each candidate $A\_{i}$ in scenario $S\_{j}$ as in (7).

(7) $V\_{total}(A\_{i},S\_{j}) = \sum\_{x=1}^{4}(W\_{x}\* V\_{x})$

From this violation score we deduce the predicted probability of the candidate by first dividing it by the sum of the violations for all $n$ candidates (in our case, $n=2)$. Because the resulting proportion represents the likelihood that the candidate will *not* be chosen, we then subtract it from 1 to obtain the probability of occurrence, as in (8).

(8) $P\_{predicted}\left(A\_{i}, S\_{j}\right)=1- ^{V\left(A\_{i}\right)}/\_{V(A\_{1})+V(A\_{2})}$

The result of this function is the predicted proportion of occurrences of the current candidate in the given scenario.

 For each of the generated weight vectors, the fitness function calculates the difference between this predicted proportion and the observed proportion of the candidate in the experimental (training) data as in (9).

(9) $Diff\left(A\_{i}, S\_{j}\right)= P\_{observed}\left(A\_{i}, S\_{j}\right)-P\_{predicted}\left(A\_{i}, S\_{j}\right) $

The mean of this difference for all $m$ scenarios in the training set, subtracted from 1 to make sure that smaller differences cause higher scores, is the fitness of the weight vector for the complete training set, as shown in (10):

(10) $Fitness \left(W\right)=1- ^{\sum\_{y=1}^{m}\sum\_{i=1}^{n}Diff(A\_{i}, S\_{y})}/\_{m}$

By selecting weight vectors that lead to a better fitness over several runs, the outcome of the genetic algorithm converges towards the optimal constraint weights for the training data.

We used the following parameters for the genetic algorithm: the population size was set to 100 and the strategy to tournamentUniform with a crossover of 0.95 and a mutation rate of 0.05. Only 5 generations were needed for the weights to converge.

## Evaluation

In the training phase, the genetic algorithm determined the optimal constraint weights for all 10 folds in the cross-validation. These values were averaged for each constraint to determine the overall weight values. For the no context conditions (scenarios 1 and 2), the optimal weights as given in (11) were obtained.

(11) \*Post/Subj 65

 \*Post/Top 48

 \*Pre/New 0

 FirstThingsFirst 0

Since in these scenarios linguistic context was absent, the constraints \*Pre/New and FirstThingsFirst are vacuously satisfied: They are not violated by any candidate, because the condition for them to have an effect (presence of context) has not been met. Therefore, they do not contribute to the violation score and hence do not receive weights.

 For the context conditions (scenarios 3 to 6) the (averaged) optimal weights are given in (12).

(12) \*Post/Subj 95

 \*Post/Top 9

 \*Pre/New 20

 FirstThingsFirst 41

Here, the constraints \*Pre/New and FirstThingsFirst do contribute to the violation score. In addition, the weight for \*Post/Top has decreased dramatically with respect to the no context conditions.

 The constraint rankings that are implied by the weight values (see also Tableaux 1 to 6) are in accordance with the differences in strength between the word order preferences that were observed in the experimental data. The subject-first preference is the strongest constraint. When it is not influenced by context, it only receives competition from the preference to put a good topic in the preverbal position. When context is present, the preference to start with the most important (i.e. new) information dominates over the preference to maintain a given-before-new structure. At the same time, the importance of prototypical topic properties is greatly reduced.

 The results that we obtained in terms of model quality (both on training data and on test data) are given in Table 3. These are again averages of the results for each of the 10 folds, with standard deviations given in brackets.

(Table 3 about here)

To test whether the prediction accuracy on the test data was significantly better than a naïve model (i.e. a model that has not been trained on the data), we validated a model that assigned random weights to the constraints in each run. We conducted a 2-sided Wilcoxon signed-rank test on the accuracy differences between the trained model and the baseline model, for each fold. The results indicate that in the no-context conditions as well as in the context conditions, the trained model performed significantly better than the naïve model (*W* = 45; *z* = 2.27; *p* < .05 and *W* = 55; *z* = 2.78; *p* < .01, respectively). This means that with the constraint weights in (12), our model is able to accurately predict the placement of bare and definite plural subjects.

 The predicted and observed proportions of the candidate NP-V-PP in each of the scenarios are presented in Table 4. Note that the proportion of the other candidate (PP-V-NP) is always one minus the proportion of NP-V-PP. The numbers indicate that using the four constraints presented above, our model is able to make accurate predictions about the likelihood that a certain word order will be chosen given a number of properties of the sentence constituents. The inclusion of weight values for the constraints is crucial here, since a traditional OT analysis would predict a 100% occurrence of the NP-V-PP order in all cases.

(Table 4 about here)

(Tableaux 1 to 6 about here)

# General discussion

The first question we tried to answer in this paper was whether a lack of prototypical topic properties influences subject position in Dutch, due to a conflict between the subject-firstpreference and the topic-firstpreference. We predicted that when no context is present, speakers will use the topic properties of the entities to choose a certain constituent order. When a lack of topic properties is paired with a prominent syntactic function such as the subject, word order choices will show variation. This prediction was borne out by the results of Experiment 1, focusing on bare plural subjects as typically non-topical constituents: A bare plural subject was put in postverbal position almost as frequently as in the canonical preverbal subject position. This is in contrast to definite subjects, which are typically good topics, and were placed in the preferred preverbal position in the majority of the cases.

Our second question concerned the interaction of prototypical topic properties with discourse requirements. We predicted that in a context that imposes a certain information structure on the sentence, the effect of topic properties is overruled. This prediction was tested in Experiment 2. The results showed that new information was either placed in preverbal position (focus preposing) or in postverbal position (given before new). However, there was no difference between the placement of bare and definite subjects, suggesting that specificity as a topic property is largely overruled by discourse properties.

 Thirdly, we asked whether we could predict the frequencies with which each word order was chosen in the experiments by modeling the preferences emerging from the data as violable psycholinguistic constraints. Given the results of our analysis, this question can be answered positively. We have shown that using an Optimality Theory framework allowed us to implement cognitive principles that are known to be important for sentence formulation as violable and possibly conflicting constraints. In addition, while traditional OT only predicts the occurrence of optimal candidates, we were able to explain the occurrence of suboptimal constituent orders using weighted constraints.

 The constraint ranking as determined by the constraint weights found by the genetic algorithm matches the observed strength of the constraints in choosing a constituent order. The fact that the NP-V-PP order was always more frequent than the PP-V-NP order, independently of context and subject type, is reflected by the high ranking of \*Post/Subj, which penalizes postverbal subjects, both in the context and in the no context conditions. This corroborates earlier findings that the subject-first principle is a strong and robust constraint on Dutch sentence formulation, as it is in many other languages (Comrie 1989; Lamers and De Hoop 2007).

 The fact that bare subjects occurred more often in postverbal position than definite subjects in Experiment 1 is reflected by a violation of \*Post/Top, which penalizes postverbal topical entities. Although the NP-V-PP order is optimal, the reverse order is relatively frequent, as indicated by the large weight value of this constraint. The fact that \*Post/Top is also violated when a topical locative PP is not in the prominent preverbal position explains why even definite subjects are sometimes produced postverbally, although they are prototypically good topics. However, the strength, and hence the effect on the production process, of the constraint \*Post/Top is susceptible to influences of the discourse context: When context is added to the sentence, its weight drops in favor of two discourse-related constraints. Here, the ranking of FirstThingsFirst above \*Pre/New matches the finding that most participants employed a strategy in which they put the most important (i.e. new) information in the sentence-initial position, while a minority adhered to the iconic given-new ordering. Note, however, that in the context model, \*Post/Top still has a weight value greater than 0. This indicates that this constraint still has some explanatory value, be it greatly diminished, when discourse context imposes a certain information structure on the sentence. Thus, our results suggest that in the presence of a constraining context prototypical topic properties may still play a role in sentence production. If, however, the effects of the presence of a discourse constraining factor are very strong, they may completely overrule the influence of prototypical topic properties. The magnitude of these effects will depend on the type of context manipulation applied (see Prat-Sala and Branigan 2000), and discourse contexts that are less strong than wh-focus questions may leave more room for effects of prototypical topic properties.

 Our model also makes new predictions about word order tendencies that were not present in the experimental data. For example, the model predicts that in a situation in which there is no constituent in the sentence that has prototypical topic properties and in which there are no discourse requirements, the preverbal position would always be occupied by the subject. It also predicts, contra Gundel (1988), that a constituent expressing important and given information may still sometimes end up in postverbal position if it lacks prototypical topic properties (i.e. specific reference). We will leave it to future research to investigate whether these predictions are borne out by real-world data.

 The present study has implications for psycholinguistic theories of language production. The results of our experiments and our analysis match an accessibility account of sentence formulation. According to this account, information that is highly accessible in memory is processed faster and tends to be produced earlier in the sentence. An entity can be accessible through its semantic properties (e.g. animate, specific), which are independent from the structure of the discourse. Accessibility can also be derived from the linguistic or non-linguistic context (e.g. given information). Both types of accessibility have been found to influence the order of constituents in a sentence (e.g. Brunetti 2009; Prat-Sala and Branigan 2000; Sridhar 1988; Van Nice and Dietrich 2003).

 In our account, entities that are highly accessible are likely to be selected as topics (Brunetti 2009; Givón 1976). In languages with a relatively free word order, linearization is often driven by information structure. In Dutch, topics can be fronted to the preverbal position in main clauses. Since entities with good topic properties are more likely to be selected as topics, they are more likely to occupy the preverbal position. However, there is also a strong preference to start a sentence with the subject. If a subject is a good topic, both grammatical function and the information structure route lead to the same word order. If this is not the case, as with subjects that are not good topics (i.e. bare plurals), the two routes may compete. Similarly, if a subject is a good topic based on the preceding context, but a less prototypical topic based on its semantic properties, inherent and derived accessibility may compete in choosing the best topic. Thus, our account is compatible with a moderately incremental view of sentence production (e.g. Ferreira and Swets 2002), in which accessibility influences the order in which entities are processed for production, but in which some planning of grammatical structure and information structure is present.

 Our account is also compatible with a model of sentence production in which many different factors are at work simultaneously and interactively to map conceptual entities onto linguistic structure in a probabilistic manner (e.g. Bates and MacWhinney 1989). Using weighted violable constraints, our model is able to predict a frequency distribution of linguistic forms, rather than viewing the occurrence of a certain structure as the outcome of the sum of a number of factors in a deterministic way.

# Conclusion

In this paper, we have presented an Optimality Theory analysis that accounts for word order variation in Dutch intransitive main clauses with bare and definite plural subjects. We used violable weighted constraints that account for the observation that Dutch main clauses may either start with the subject or with the topic, and that what constitutes a good topic is influenced by both prototypical topic properties and discourse requirements. In this way, we were able to accurately predict the frequencies of occurrence of two competing word orders obtained from two sentence production experiments.

 As typically bad topics, bare plural subjects are not very suitable candidates to be placed in the preverbal position. As subjects, on the other hand, they are preferred in this position. In addition, discourse factors may impose a certain information structure on a sentence, which may also affect word order. By taking into account both grammatical function and factors that tap on information structure characteristics, we can straightforwardly account for the word order variation in sentences with bare plural subjects. Moreover, by separating constraints that relate to inherent accessibility from those relating to derived accessibility in our analyses, we can also explain the finding that the influence of prototypical topic properties disappears to satisfy certain discourse requirements as indicated by the presence of the discourse context. Of course, other factors than those covered here may play a role. Recent corpus data (Vogels and Van Bergen submitted), for instance, suggests that bare subjects higher in animacy are more likely to occupy the preverbal position. Such effects could be a reason to include additional constraints in the model. We leave this to future research.

 As we have shown, our analysis is compatible with a psycholinguistic account in which word order variation is explained in terms of accessibility. By combining psycholinguistic production models with a computational approach in an OT framework, our analysis contributes to bridging the gap between psycholinguistics on the one hand and computational and theoretical linguistics on the other. With this interdisciplinary approach we were able to show that a combination of grammatical, conceptual and discourse factors probably serves best to explain word order variation in Dutch clauses with bare plural subjects.

**Acknowledgements**

This research was partly supported by the Netherlands Organization for Scientific Research (NWO) (grant 360-70-220, Animacy). Part of this research was presented as a poster at the 6th International Workshop on Language Production in Edinburgh, 2-4 September 2010. We thank the audience, in particular Christoph Scheepers, for their useful comments. We also thank Helen de Hoop, Emiel Krahmer and Piek Vossen for their comments on earlier versions of this paper. Special thanks go to Kees de Schepper for programming the drag-and-drop experiments.

References

Aissen, J., 1999. Markedness and Subject Choice in Optimality Theory. Natural Language & Linguistic Theory 17, 673-711.

Aissen, J., 2003. Differential Object Marking: Iconicity vs. Economy. Natural Language and Linguistic Theory 21, 435-483.

Bates, E., MacWhinney, B., 1989. Functionalism and the Competition Model, in: Bates, E., MacWhinney, B. (Eds.), The Crosslinguistic Study of Sentence Processing. Cambridge University Press, Cambridge, UK, pp. 3-73.

Boersma, P., Hayes, B., 2001. Empirical tests of the gradual learning algorithm. Linguistic Inquiry 32, 45-86.

Bouma, G.J., 2008. Starting a sentence in Dutch: a corpus study of subject- and object-fronting. Rijksuniversiteit Groningen.

Branigan, H.P., Pickering, M.J., Tanaka, M., 2008. Contributions of animacy to grammatical function assignment and word order during production. Lingua 118, 172-189.

Brunetti, L., 2009. On the semantic and contextual factors that determine topic selection in Italian and Spanish. The Linguistic Review 26, 261-289.

Chafe, W.L., 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view., in: Li, C.N. (Ed.), Subject and topic. Academic Press, New York:, pp. 25-56.

Comrie, B., 1989. Language universals and linguistic typology, 2nd ed. University of Chicago Press, Chicago.

Crasborn, O., van der Kooij, E., Ros, J., de Hoop, H., 2009. Topic agreement in NGT (Sign Language of the Netherlands). The Linguistic Review 26, 355-370.

Ferreira, F., Swets, B., 2002. How Incremental is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. Journal of Memory and Language 46, 57-84.

Givón, T., 1976. Topic, Pronoun, and Grammatical Agreement, in: Li, C.N. (Ed.), Subject and Topic. Academic Press, New York/San Francisco/London, pp. 149-188.

Givón, T., 1988. The pragmatics of word order: predictability, importance and attention, in: Hammond, M., Moravcsik, E.A., Wirth, J.R. (Eds.), Studies in Syntactic Typology. John Benjamins, Amsterdam/Philadelphia.

Goldberg, D.E., Holland, J.H., 1988. Genetic algorithms and machine learning. Machine Learning 3, 95-99.

Goldwater, S., Johnson, M., 2003. Learning OT constraint rankings using a maximum entropy model, in: Spenader, J., Eriksson, A., Dahl, Ö. (Eds.), Proceedings of the Stockholm Workshop on Variation within Optimality Theory. Stockholm University, pp. 111–120.

Gundel, J.K., 1988. Universals of topic-comment structure, in: Hammond, M., Moravcsik, E.A., Wirth, J.R. (Eds.), Studies in Syntactic Typology. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Haiman, J., 1983. Iconic and Economic Motivation. Language 59, 781-819.

Jäger, G., 2007. Maximum Entropy Models and Stochastic Optimality Theory, in: Zaenen, A., Simpson, J., King, T.H., Grimshaw, J., Maling, J., Manning, C. (Eds.), Architectures, Rules, and Preferences. Variations on Themes by Joan W. Bresnan. CSLI Publications, Stanford, pp. 467-479.

Keenan, E.L., Comrie, B., 1977. Noun Phrase Accessibility and Universal Grammar. Linguistic Inquiry 8, 63-99.

Keller, F., 2006. Linear Optimality Theory as a model of gradience in grammar, in: Fanselow, G., Féry, C., Vogel, R., Schlesewsky, M. (Eds.), Gradience in grammar: Generative perspectives. Oxford University Press, Oxford, pp. 270–287.

Kempen, G., Hoenkamp, E., 1987. An Incremental Procedural Grammar for Sentence Formulation. Cognitive Science 11, 201-258.

Lambrecht, K., 1994. Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge University Press, New York.

Lamers, M., de Hoop, H., 2005. Animacy information in human sentence processing: An incremental optimization of interpretation approach, in: Christiansen, H., Skadhauge, P.R., Villadsen, J. (Eds.), Constraint solving and language processing. Springer-Verlag, Berlin.

Lamers, M.J.A., De Hoop, H., 2007. Object fronting in Dutch. Under review.

McDonald, J., Bock, K., Kelly, M.H., 1993. Word and World Order: Semantic, Phonological and Metrical determinants of Serial Position. Cognitive Psychology 25, 188-230.

Prat-Sala, M., Branigan, H.P., 2000. Discourse constraints on syntactic processing in language production: a cross-linguistic study in English and Spanish. Journal of Memory and Language 42, 168-182.

Prince, A., Smolensky, P., 1993. Optimality Theory: Constraint interaction in generative grammar. Ms. Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder.

Pulleyblank, D., Turkel, W.J., 1998. The logical problem of language acquisition in Optimality Theory, in: Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., Pesetsky, D. (Eds.), Is the best good enough? Optimality and competition in syntax. MIT Press, Cambridge, Mass., pp. 399–420.

Reinhart, T., 1982. Pragmatics and Linguistics: An Analysis of Sentence Topics. Philosophica 27, 53-94.

Smolensky, P., Legendre, G., 2006. The harmonic mind: From neural computation to optimality-theoretic grammar. . MIT Press, Cambridge, MA, US.

Sridhar, S.N., 1988. Cognition and sentence production: a cross-linguistic study. Springer-Verlag, New York.

Van Bergen, G., De Swart, P., 2010. Scrambling in spoken Dutch: Definiteness versus weight as determinants of word order variation. Corpus Linguistics and Linguistic Theory 6, 267-295.

Van Nice, K.Y., Dietrich, R., 2003. Animacy effects in language production: From mental model to formulator, in: Härtl, H., Tappe, H. (Eds.), Mediating between Concepts and Grammar. Mouton de Gruyter, Berlin/New York, pp. 101-117.

Van Tiel, B., Lamers, M.J.A., 2007. Animacy in verschillende teksttypes. Tabu: Bulletin voor Taalwetenschap 36, 19-38.

Vogels, J., van Bergen, G., submitted. Where to place the subject in Dutch: The role of prototypical topic properties.

Ward, G., 1985. The semantics and pragmatics of preposing. University of Pennsylvania.

Yang, N., Van Bergen, G., 2007. Scrambled Objects and Case Marking in Chinese. Lingua 117, 1617-1635.

1. http://www.neurobs.com [↑](#footnote-ref-1)
2. One participant was excluded from the analysis, because this participant showed a significant overall bias for placing the top constituent in the left slot. In addition, there were 5 unscorable experimental items due to exceeding the time limit. Allresponse items but one were answered correctly, indicating that all participants paid attention to the task. [↑](#footnote-ref-2)
3. No inanimate NPs were included in the experimental material, because this would lead to the inclusion of what-questions, which in Dutch, as in English, are awkward with plural NPs. [↑](#footnote-ref-3)
4. One trial with a reaction time above 3 standard deviations from the overall mean was removed. Thirty trials with reaction times above 2.5 standard deviations from the participant’s mean were replaced by the mean. Given that there was only 1 error in the response items, participants were assumed to have paid attention to the task. [↑](#footnote-ref-4)
5. The constraints \*Pre/New and FirstThingsFirst correspond to Gundel’s (1988) *Given Before NewPrinciple* and the *First Things First Principle*, which she used to explain the cross-linguistic distribution of topic-focus constructions. [↑](#footnote-ref-5)
6. We used the implementation of a Genetic Algorithm in Perl (<http://search.cpan.org/~aqumsieh/AI-Genetic-0.04/Genetic.pm>). [↑](#footnote-ref-6)
7. In this experiment, we only distinguish two alternatives but the same method can be applied to a larger number of alternatives. [↑](#footnote-ref-7)