Cognition 254 (2025) 105999

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Semantic interference across word classes during lexical selection in Dutch

Constantijn L. van der Burght [a,b,*], Antje S. Meyer [a,c]

[a] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[b] Leiden University Centre for Linguistics, Leiden University, Leiden, the Netherlands
[c] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Language
Sentence production
Semantic interference
Lexical access
Syntax
Picture naming

ABSTRACT

When producing a sentence, speakers must rapidly select appropriate words in the correct order. Models of lexical access often assume that this lexical selection process is competitive and that each word is chosen from a set of competing candidates. Therefore, an important theoretical issue is which factors constrain this choice. Speech error evidence suggests that word class plays a decisive role here and that lexical access is, at any point in time, restricted to words that fit the part of the grammatical structure of the sentence that is being constructed. Using a novel version of the picture-word interference paradigm, Momma, Buffinton, Slevc, and Phillips (2020, Cognition) showed experimentally that word class indeed constrains lexical selection. Specifically, in speakers of American English, action verbs (as in *she's singing*) competed with semantically related action verbs (as in *she's whistling*), but not with semantically related action nouns (as in *her whistling*). Similarly, action nouns only competed with semantically related action nouns, but not with action verbs. As this pattern has important implications for models of lexical access and sentence generation, we conducted a conceptual replication of the study in Dutch. In two experiments, we found a semantic interference effect, but, contrary to the original study, no evidence for a word class constraint. In accounting for these results, we propose that word class constraints on lexical selection are graded rather than categorical, and that, at least for verbs and action nouns, the marking for word class is clearer in English than in Dutch.

1. Introduction

A central issue for theories of sentence generation is how speakers manage to produce words in the appropriate order and at high speed, often at rates of three or four words per second. Selecting the right word at the right time is challenging because adults have a mental lexicon of tens of thousands of items to select from. Thus, when a speaker plans to say a sentence, e.g., “the cat sits on the roof”, one might expect competition between candidate words for each sentence position (e.g., between “cat” and “kitten”) and between candidate words for different positions (e.g., “cat” and “sit”), which need to be produced in quick succession. A common assumption in models of sentence generation is that word order rules are implemented in such a way that they constrain the set of lexical items competing for selection. At any moment during the generation of an utterance, only those lexical items that fit into the grammatical structure of the developing utterance plan are considered for selection. Such a word class (or syntactic category) constraint explains an important observation about speech errors: Healthy adults rarely produce word exchange errors and, when such errors do occur, they almost always involve words of the same word class (Fromkin, 1971; Nooteboom, 1973). A word class constraint on lexical selection was originally proposed to explain this important observation. Additionally, a possible word class constraint plays a critical role in current theories of sentence production, where it is implemented as a mechanism that biases word selection towards candidates that fit into the developing sentence structure. These biases reduce the competition between words in different sentence positions and support the selection of words that fit into the grammatical structure of the utterance.

Because of the proposed pivotal role of the word class constraint in sentence planning, it is important to assess whether lexical selection is indeed constrained by word class. While word exchange errors are constrained by word class, such errors are infrequent, and word class is confounded with conceptual features, so that it is difficult to estimate the strength of a purely grammatical constraint. The results of experimental work aiming to establish the impact of a word class constraint on lexical selection are mixed and often difficult to interpret. However, a study by Momma et al. (2020) provided compelling evidence for a word class constraint on lexical selection. Given the important implications of

* Corresponding author at: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands.
E-mail address: stan.vanderburght@mpi.nl (C.L. van der Burght).

https://doi.org/10.1016/j.cognition.2024.105999
Received 6 March 2024; Received in revised form 21 October 2024; Accepted 22 October 2024

these results for theories of sentence generation, we conducted a conceptual replication of this study. Before turning to the study by Momma and colleagues and our own experiments, we briefly review relevant earlier studies.

The studies reviewed below used versions of the picture-word-interference paradigm, where participants name or describe pictures accompanied by written or spoken distractor words. A key finding in this paradigm is the semantic interference effect (SIE; see Bürki et al. (2020) for a review): Participants are slower to name target pictures (e.g., a picture of a cat) accompanied by written or spoken distractor words that belong to the same semantic category as the target (e.g., dog) than target pictures accompanied by unrelated distractor words (e.g., spoon; Lupker, 1979; Schriefers et al., 1990). The SIE suggests that lexical access occurs *competitively*: When a picture name is accessed, semantically related names are activated as well, creating competition among them, which needs to be resolved (Levelt et al., 1999; Rahman & Melinger, 2019; Roelofs, 1992; for an alternative view, discussed in the General Discussion, see Finkbeiner & Caramazza, 2006; Mahon et al., 2007). More specifically, when a target and distractor are conceptually related, there is mutual activation of the two concepts, which spreads to the associated grammatical word representations (the lemmas) and, subsequently, to the morpho-phonological representations. The distractor benefits more from this co-activation than the target because the target is more highly activated than the distractor, and therefore sends more activation to the distractor than it receives from it. This is because the speaker attends to the target and aims to ignore the distractor. No mutual activation occurs when target and distractor are semantically unrelated. Therefore, stronger competition occurs between the target and a related than an unrelated distractor concept and lemma. As this competition needs to be resolved before the target lemma can be selected, a delay occurs in lexical selection and picture naming.

Picture naming latencies in the presence of different types of distractors reveal how strongly different types of words interfere with the selection of target lemmas. Therefore, the picture-word interference paradigm can be used to assess whether lexical selection is constrained by word class. In an early study using this approach with speakers of German, Pechmann and Zerbst (2002) see also Pechmann et al., 2004) found longer object naming latencies when distractors were nouns, like the picture names the participants produced, rather than belonging to other word classes. However, results of a replication in German and a follow-up study in English conducted by Janssen et al. (2009) suggested that the word class effect seen in these studies was semantic rather than syntactic in nature, driven by a difference in imageability of the distractors. In Vigliocco et al. (2005), speakers of Italian named action pictures in infinitive or inflected verb forms in the presence of semantically related or unrelated distractor nouns or verbs (in infinitive form). The authors found an SIE regardless of the word class of the distractors. In addition, there was an effect of word class, with longer naming latencies after verb than noun distractors. Importantly, this main effect of word class only arose when the participants produced the inflected form (e.g., *she cries*) rather than the infinite form of the verb (e.g., *to cry*). The authors argued that this pattern arose because sentence integration was only required in the inflection condition. However, the word class effect was not replicated in a study with Japanese speakers (Iwasaki et al., 2008). Instead, these experiments showed semantic interference effects independent of the word class of the distractor. In a study with English-speaking participants, Mahon et al. (2007) used object naming and noun and verb distractors, and found an interaction between word class congruency and semantic relatedness. However, based on the results of a control experiment, they attributed the prolonged latencies after related noun compared to related verb distractors to a difference in imageability of the distractor types. Finally, Rodríguez-Ferreiro et al. (2014) studied both object and action naming in speakers of Spanish. In two experiments, they found interactions between semantic relatedness and word class congruency: Semantic interference was only observed when target and distractor belonged to the same word class (verb or noun) but not when they belonged to different word classes. However, as in the experiments mentioned above, different lexical items were used as noun and verb distractors (e.g., translation equivalents of *bone* and *to bite* for the target *to bark*), and so related noun and verb distractors may have differed in their similarity to the targets. The authors addressed this issue in a third action naming experiment, where, in addition to object nouns (*bone*) they also included action nouns (*bite*, morphologically marked as noun) and verbs (*bite*, morphologically marked as verb). Participants produced either citation or inflected forms of the target verbs. As expected, an SIE effect was found from verb distractors but not object noun distractors. Interestingly, however, an SIE from action noun distractors only arose when participants produced citation forms of verbs, but not when they used the inflected forms. This pattern contrasts with the results obtained from Italian speakers by Vigliocco et al. (2005), who found an interference effect of verb compared to noun distractors when target verbs were produced in the inflected, but not citation forms.

A recurrent challenge in this body of research has been to vary word class congruency between targets and distractors without simultaneously varying critical properties of the distractors, such as their imageability, and/or the semantic similarity between target and distractor. Rodríguez-Ferreiro et al. (2014) addressed this issue in their third experiment, but the results were inconclusive. Momma et al. (2020) also used matched action verbs and nouns as distractors, and, different from Rodriguez and colleagues, also used both word classes as targets. As this study motivated our own work, we describe it in some detail.

The study by Momma et al. consisted of two experiments. In both of them, native speakers of English read sentences where the sentence-final distractor word was either an action noun *(They were surprised by her beautiful singing*) or a verb *(They were surprised that she was beautifully singing*). In Experiment 1, the participants then described a picture using a verb that was semantically related or unrelated to the distractor word (*she's whistling* or *she's baking*). Filler trials in which participants had to reproduce the written sentences ensured that the distractors were accurately processed. This paradigm differs from the classic picture word interference paradigm because the distractor words were not presented simultaneously with the target words. Instead, distractor words had already been accessed and had to be held in working memory. Nevertheless, semantic interference was found. Importantly, semantic interference only arose when the distractors were verbs, like the targets. In Experiment 2, the same sentences were used, but the participants produced action nouns (*Her singing* …) instead of verbs. Here, semantic interference only arose when the distractors were action nouns as well. The authors concluded that word class was a tight constraint on lexical selection.

Momma and colleagues interpret their findings within the spreading activation model proposed by Dell et al. (2008), see also Gordon & Dell, 2003, and Chang et al., 2006), dubbed the *syntactic traffic cop* model. Consistent with other models of lexical organisation and sentence generation (e.g., Bock & Levelt, 1994; Jackendoff, 2003), Dell and colleagues assume that during sentence generation, lexical units become activated and selected. Lexical units are specified for semantic features as well as word class. In addition to these lexical units, there are syntactic sequential state units which capture word order rules, such as *determiner-noun-verb* for sentences such as *the bird flies.* While a message is encoded, the conceptual input remains static, but the activation of the syntactic state units changes. These units sequentially activate matching lexical units and inhibit mismatching ones. For instance, when the speaker plans the utterance *the bird flies* and has just selected the determiner, the syntactic state nodes boost the activation of all nouns and reduce the activation of determiners and verbs. In this way, the architecture establishes dynamically changing biases towards units that fit the grammatical structure of the developing utterance plan. It explains the word class constraint on lexical errors as well as the word class constraint in picture-word interference experiments.

The design of the study by Momma and colleagues stands out from earlier work because distractor word class and semantic relatedness to the target were manipulated independently from one another. Furthermore, both action nouns and verbs appeared as distractors as well as targets. As the same words were used across both word class conditions, the distractors were matched in phonological forms as well as for semantic and conceptual variables such as imageability. Furthermore, semantic relatedness between distractors and targets was tightly controlled. The interaction between word class congruency and semantic relatedness suggests that both effects arise at the same level of processing, most likely, in our view, during the selection of lexical items. Contrary to the additive effect of word class and semantic relatedness found by Vigliocco et al. (2005), the interaction suggests a tight word class constraint on lexical selection. Apparently, only distractors of the target-appropriate word class entered the competition for selection. The results therefore suggest that the distractor word class was pre-specified, for instance in the form of a syntactic frame, prior to lexical selection.

In their recent review and meta-analysis of the results of picture-word interference studies, Bürki et al. (2020) demonstrated the robustness of the semantic interference effect itself, yet also highlighted that results concerning variables that may influence the effect (e.g., semantic distance between target and distractor or distractor frequency) require replication. Considering the broad theoretical implications of the results obtained by Momma et al. (2020), we aimed to replicate the results in a new study with speakers of Dutch. The rationale, hypotheses, and analysis plan of the study were pre-registered at the Open Science Framework. For our study, we converted the paradigm from the English original to Dutch, following the original design as closely as possible. Dutch, like English, allows for the use of nominalised verbs *(het zingen, the singing*). In Dutch, these action nouns are homophonous to the third person plural present conjugation of the verb (*zij zingen, they sing*) and the infinitive (*zingen, to sing*). As in the original study, the two forms were disambiguated by the syntactic structure of the distractor sentence. Furthermore, the required target sentence structure ensured that the target word was produced as a verb (*zij zingen*, *they sing*; Experiment 1) or noun (*hun zingen is begonnen*, *their singing has begun;* Experiment 2). As in the original study, target and distractor were either semantically related or unrelated. For both experiments, we expected an interaction between word class congruency and semantic relatedness of targets and distractors. In Experiment 1, the SIE should only occur after verb distractors, whereas in Experiment 2 it should only occur after noun distractors.

The current study was conducted in Dutch primarily for practical reasons, namely easy access to a Dutch-speaking participant sample. In addition, our study allowed us to assess whether the results obtained by Momma and colleagues would be replicated in a different language. The semantic interference effect has been amply documented for English and Dutch (Costa et al., 2005; Mahon et al., 2007; Roelofs, 1992; Schriefers et al., 1990; see also Bürki et al., 2020), but no earlier study with Dutch speakers had examined whether this effect would interact with word class congruency. We assumed that the general processing mechanisms involved in the interplay between lexical access and grammatical encoding would be the same for English and Dutch. Therefore, a processing model such as the syntactic traffic cop model proposed by Dell and colleagues should apply similarly to both languages. Consequently, we expected to replicate the interaction between semantic relatedness and word class congruency observed by Momma and colleagues.

However, as reviewed above, earlier studies of syntactic constraints on lexical access have yielded inconsistent results, and, as these studies were carried out in different languages, cross-linguistic differences may have contributed to the heterogeneity of the findings. Dutch and English are structurally highly comparable (Barbiers, 2000; van Haeringen, 1956), though word order may be somewhat freer in Dutch than in English (Koster, 1999). Moreover, nominalizations are derived in different ways from the stem, and so the corresponding morphological representations differ across the two languages. These and other differences between the languages may affect the lexical access process, but we had no specific predictions about the way they might affect the results of the experiments. The prediction for both experiments was that the interaction observed by Momma and colleagues should be replicated.

2. Methods

The theoretical motivation, hypotheses, experimental design and analysis plan of this study were preregistered at the Open Science Framework (https://osf.io/gy82u/). The design and analysis plan followed the original study (Momma et al., 2020). In designing and analysing the experiments, we proceeded as proposed in the pre-registration document. Any deviations from the proposal are mentioned.

2.1. Participants

60 participants took part in Experiment 1 (40 female, 20 male; age: $M = 22.8$ years, $SD = 2.8$, range = 18–30). In Experiment 2, 60 participants took part who had not participated previously (46 female, 14 male; age: $M = 22.0$ years, $SD = 3.2$, range = 18–32). Power analyses (*simr* in R, version 1.0.6) performed on the results obtained by Momma et al. suggested that to achieve a power of .80, sample sizes of 60 (Experiment 1) and 48 (Experiment 2) were required. We chose to have equal sample sizes in both experiments, and so sampling continued until 60 complete datasets were acquired. Participants were native speakers of Dutch reporting no hearing or language disorders, recruited from the database of the Max Planck Institute for Psycholinguistics. All were paid 15€ for participation and gave informed consent prior to the experiment. The study was approved by the Ethics Committee Faculty of Social Sciences, Radboud University, Netherlands.

2.2. Materials

The experimental design followed the study by Momma et al. (2020) as closely as possible. A picture-word interference paradigm was used involving distractor sentences rather than single words. The design involved the within-participant factors *Distractor relatedness*, manipulating the semantic relationship between distractor and target (related; unrelated) and *Distractor word class*, defined by the syntactic (grammatical) category of the sentence-final word of the distractor sentence (verb; noun). In Experiment 1 targets were realised as verbs and in Experiment 2 as nouns (see Table 1 for examples).

The targets were pictures depicting actions that corresponded to (optionally) intransitive verbs (e.g., zingen / *to sing*, koken / *to cook*, fluiten / *to whistle* etc.). We used the 24 target pictures from the original study (developed by Szekely et al. (2004)) except for *hoesten* (*coughing*) and *niezen* (*sneezing),* which we replaced by new pictures displaying the actions more clearly. Furthermore, as our experimental design required verbs to be uttered in third person plural form, we showed two identical actors in each picture (see Fig. 1).

The target words were paired with distractor words to create semantically related combinations (e.g., *whispering–calling*). The words were re-combined to create semantically unrelated pairs (e.g., *walking–calling*). As a result, all target pictures were seen four times, combined with related and unrelated noun distractors, and with related and unrelated verb distractor sentences. The carrier sentences (*Marie vertelde dat…/ Marie told that.*) were similar for each noun/verb pair in each related or unrelated pairing with the target (see Table 1). The pairing of these sentences (to related or unrelated distractors) was counterbalanced across participants. The semantic relatedness between distractor-target pairs was measured with a latent semantic analysis (Landauer & Dumais, 1997) using a semantic space pre-trained on a Dutch web-based corpus (Grave et al., 2018). The cosine distance of the related pairs ($M = 0.57$; $SD = 0.10$) and the unrelated pairs ($M = 0.34$; $SD = 0.07$) was significantly different (two-tailed *t*-test; $t(23) = 14.2$, $p < .001$).

Table 1
Overview with example distractor and target sentences (translations in English in italics).

Distractor relatedness	Distractor word class	Distractor sentence	Target sentence Experiment 1	Target sentence Experiment 2
Related	Verb	Zij is ervan onder de indruk dat wij mooi zingen *She is impressed that we beautifully sing*	Zij fluiten *They whistle*	Hun fluiten is begonnen/afgelopen *Their whistling has begun/ended*
	Noun	Zij is onder de indruk van ons mooie zingen *She is impressed by our beautiful singing*		
Unrelated	Verb	Jan vertelde de dokter dat jullie aanhoudend hoesten *John told the doctor that you persistently cough*		
	Noun	Jan vertelde de dokter over hun aanhoudende hoesten *John told the doctor about their persistent coughing*		

2.3. Procedure

Each experiment started with a familiarisation phase in which participants received a booklet with target pictures and the corresponding verbs presented in the citation form (e.g., *zingen*). Participants were given as much time as they required to study the pictures and words. During the experiment, participants sat in a sound-attenuated booth in front of a computer monitor and microphone. Each session started with practice trials containing two distractor-target pairs that were not included in the actual experiment. The instructions were to carefully read and memorise the sentence on the screen. Upon reading *Herhaal* (*repeat*), participants should repeat the sentence. Upon seeing a target picture, they should describe the picture as quickly and fluently as possible, avoiding hesitations. In contrast to traditional picture-word interference paradigms, the task involved distractor sentences as opposed to single words. Moreover, in the current study, participants were instructed to remember the distractors, whereas in traditional versions of the paradigm the distractors should be ignored.

In Experiment 1, participants used the personal pronoun *zij* (*they*) followed by the target verb, yielding declarative sentences such as *zij fluiten* (*they whistle*). In Experiment 2, they described that the action had started or ended, as cued by a green or red clock, respectively (see Fig. 1). In this experiment, the participants' utterances started with the possessive pronoun *hun* (*their*) followed by the target noun and either *is begonnen* (*has begun*) or *is afgelopen* (*has ended*), as in *hun fluiten is begonnen* (*their whistling has begun)*. The use of *begin* and *end* constituted a departure from the utterances elicited in the original study. Momma et al. (2020) employed a cover story where participants had to imagine themselves as a person with synaesthesia, perceiving certain actions as having colours. The target pictures were presented with a small coloured square in the corner of the picture, prompting participants to produce sentences such as *Her whistling is red* or *Her whistling is green,* depending on the colour of the square. To simplify the task, we used target sentences such as *Their whistling has begun/ended.* In line with the original study, a colour cue ensured that participants could produce the desired structure with minimal visual processing. Crucially, both versions of the task lead participants to produce nominalised verb forms (*fluiten*, *whispering,* as noun).

Experimental trials started with a fixation cross lasting for 800 ms. Subsequently, the distractor sentence was presented on the screen. Upon button press, either the word *Herhaal* (*repeat*; 50 % of trials) or the target picture (50 % of trials) was presented. Presentation continued until participants pressed *enter*. After a 1000 ms delay the next trial started. Visual stimuli were presented on a screen (LCD monitor GN246HL, Acer) and button responses were given via a response box placed in front of the participant. Stimulus presentation and speech recordings were

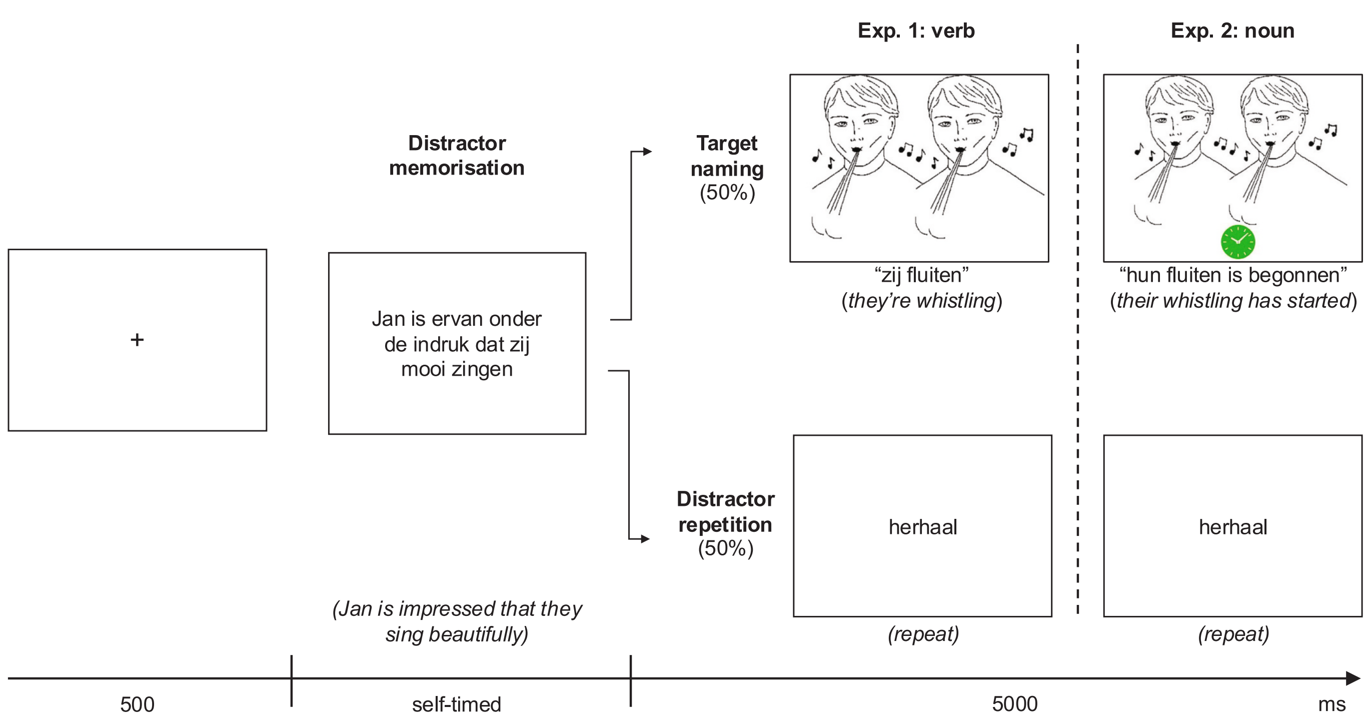

Fig. 1. Overview of the experimental trial. Each trial started with a fixation cross followed by the distractor sentence, which the participants were asked to memorise. Following a button press, a target picture appeared. The picture prompted the participants to produce the target word as either verb (Experiment 1) or as noun (Experiment 2). On 50 % of trials participants repeated the distractor sentence, cued by the word *herhaal* (*repeat*).

controlled using Presentation (Neurobehavioural Systems, Inc., Albany, CA, USA). The trial sequences were pseudo-randomised using the software Mix (Casteren & van Casteren & Davis, 2006) with the constraints that the same trial type (experimental or filler trials) could not be presented more than three times in succession, and the same distractor word class or relatedness condition not more than twice. Per block of 48 trials, each target picture occurred only once. The experiment lasted for around 40 min, including three self-paced breaks.

2.4. Data analysis

The data analysis adhered to the pre-registered analysis plan, which in turn corresponded to the analysis performed by Momma et al. Speech recordings were analysed in Praat (Boersma & Weenink, 2022). Speech onset was automatically detected by marking the time at which intensity exceeded 65 dB, and manually adjusted by trained annotators, who also determined the onsets of the target verb or noun. The annotators were unaware of the distractor-target pairing. We analysed both speech and target onsets since speakers may differ in their sentence planning strategies: They may plan the target before utterance onset or during the production of the pronoun. Consequently, an interference effect might arise in the sentence onset latencies, the target onset latencies, or both.

Following the original study, we excluded non-target utterances (containing wrong lexical items and/or a wrong sentence structure) as well as utterances containing filled pauses (e.g., "eh"; Exp. 1: 7.1 %; Exp. 2: 12.8 % of all trials). Subsequently, we removed trials with a speech onset latency below 300 ms or above 5000 ms (Exp. 1: 0.04 %; Exp. 2: 0 % of non-erroneous trials). In the *repeat* trials, only verbatim reproductions of the distractor sentence were considered correct.

Statistical analysis was performed with linear mixed models (LMMs) using the lme4 package (Bates et al., 2015) in R (R Core Team, 2024). For the fixed effects analysis, factors were dummy coded using a sum-to-zero contrast. The full model included fixed effects for *Distractor relatedness, Distractor word class*, and their interaction. Centred *Trial order* was included as an additional fixed effect. We aimed for a maximal random effects structure (Barr et al., 2013), simplifying this in case of convergence issues by removing the random slope that accounted for the least amount of variance until the model converged (Momma et al., 2020). In our pre-registered analysis plan, we committed to performing an LMM on log-transformed latencies, in accordance with Momma et al. In addition, in our pre-registration we planned to run a generalised LMM on raw latencies (using a gamma distribution) to avoid well-known issues with log-transformed data (Lo & Andrews, 2015). However, an evaluation of the model diagnostics on both model types favoured the models on log-transformed latencies. The added benefit of this analysis is that it is in line with the analysis reported in Momma et al. Indeed, it is considered best practice to report one confirmatory model and we therefore opted to report the results of the analysis using log-transformed latencies only. Alternative models including model diagnostics plots are available on OSF. Statistical inference was performed according to the *t*-value associated with the fixed effects parameter estimates. *T*-values greater 1.96 were considered significant.

3. Results and discussion

3.1. Experiment 1

On *Repeat* trials, participants correctly repeated the preceding distractor sentences in 92.9 % ($SD = 4.4$ %) of the trials. As only verbatim reproductions of the distractors were considered correct, this indicates that participants engaged with the task, and read, remembered, and recalled the lexical and morphological information from the distractors with high accuracy. This also indicates that performances of the Dutch speakers matched performance of the English speakers in Experiment 1 of the original study (83.3 % correct repetitions).

For the sentence onset latencies on experimental trials (for descriptive statistics see Table 2), a significant main effect of distractor relatedness was found ($\beta = 0.011$, $SE = 0.003$, $t = 3.472$): sentence onsets occurred, on average, 18 ms later after related as compared to unrelated distractor sentences. There was no significant main effect of distractor word class nor was there an interaction between these two factors (all *t*-values >1.96; see Table 3). The same pattern of results was found for the verb onsets. There was a main effect of distractor relatedness ($\beta = 0.008$, $SE = 0.003$, $t = 3.225$), with later sentence onsets, by 18 ms, for related than unrelated distractor sentences. The effect of distractor word class and the interaction effect were not significant (all *t*-values >1.96; see Table 4). Overall, onset latencies were faster for the Dutch as compared to English speakers, while the verb onsets were similar across the two studies.

In sum, the results indicated good performance in the repetition task, demonstrating attention to the task and encoding of the grammatical structure. Furthermore, the main effect of relatedness for sentence as well as verb onsets indicated full planning of the utterance before speech onset, consistent with the results by Momma et al. Additionally, the main effect indicated that the related as compared to unrelated distractors gave rise to a semantic interference effect (Fig. 2). Most importantly, this effect was not significantly modified by word class, deviating from the results found by Momma et al. In the original study, an interaction between distractor relatedness and distractor word class was found, showing a word class-specific SIE. We return to this pattern in the General Discussion.

3.2. Experiment 2

In the *Repeat* trials, participants correctly repeated 87.4 % ($SD = 6.7$ %) of the distractor sentences. This indicates that, as in Experiment 1, they were engaged with the task and processed the lexical and morphological information presented in the distractors. Again, performance of the Dutch speakers matched the performance of the English speakers in Experiment 2 of the original study (81.0 % correct repetitions).

In this experiment, speakers were instructed to produce action nouns instead of verbs. For the sentence onset latencies (for descriptive statistics see Table 5), a significant main effect of distractor relatedness was found ($\beta = 0.013$, $SE = 0.005$, $t = 2.804$): Sentence onsets were 28 ms later after related vs. unrelated distractors. There was no main effect of distractor word class, nor was there an interaction between the two terms (all *t*-values <1.96; see Table 6). This suggests that the SIE was independent from the type of distractor word class (noun or verb). The same pattern of results was found when considering the onsets of the nouns. Again, there was a significant main effect of distractor relatedness ($\beta = 0.011$, $SE = 0.004$, $t = 2.612$), with an SIE of 27 ms. The distractor word class and interaction effects were not significant (all *t* – values <1.96; see Table 7). Again, onset latencies were faster for the Dutch as compared to the English speakers, while the noun onsets were similar across the two studies.

In sum, the results of Experiment 2 showed good performance in the repetition task, indicating careful processing of the distractor sentences and their morphosyntactic properties. There was a main effect of relatedness in both dependent variables (sentence onsets and noun onsets), indicating full planning of the sentences before speech onset and interference from the related distractors. Importantly, there was no

Table 2
Results of Experiment 1: means and standard errors (*SEM*) for sentence and verb onsets after related and unrelated noun and verb distractors (all values in ms).

word class	relatedness	mean onset verb (*SEM*)	mean onset (*SEM*)
noun	related	1119 (47)	842 (43)
noun	unrelated	1103 (47)	827 (42)
verb	related	1108 (44)	835 (41)
verb	unrelated	1088 (45)	814 (41)

Table 3
Estimates of linear mixed effects model on log-transformed sentence onset latencies (Experiment 1). The model included distractor word class, distractor relatedness, their interaction, and trial order (scaled) as fixed effects. Random effects included by-participant intercepts and slopes for word class and by-item intercepts and slopes for relatedness.

term	estimate	std. error	*t*-value
intercept	6.672	0.034	195.050
relatedness	0.011	0.003	3.472
word class	0.005	0.003	1.525
trial order (scaled)	−0.004	0.000	−35.734
relatedness * word class	−0.003	0.003	−0.875

Table 4
Estimates of linear mixed effects model on log-transformed verb onset latencies (Experiment 1). The model included distractor word class, distractor relatedness, their interaction, and trial order (scaled) as fixed effects. Random effects included by-participant intercepts and slopes for word class and by-item intercepts and slopes for relatedness.

term	estimate	std. error	*t*-value
intercept	6.975	0.030	233.879
relatedness	0.008	0.003	3.225
word class	0.004	0.003	1.549
trial order (scaled)	−0.003	0.000	−36.471
relatedness * word class	−0.002	0.003	−0.630

interaction with distractor word class: there was comparable interference from verbs and nouns (Fig. 2). Again, the results do not replicate the results found in the original study: Momma et al. (2020) found word class-specific semantic interference, with only related vs. unrelated nouns (not verbs) yielding prolonged naming latencies.

4. General discussion

In a picture description task, Momma et al. (2020) found that the production of action verbs was hindered when participants had just read and planned related action verbs, compared to unrelated ones. Conversely, the production of action nouns was hindered by prior reading and planning of related compared to unrelated action nouns. There was no interference across word categories: Related action nouns did not hinder the production of action verbs, and related action verbs did not hinder the production of action nouns. These results are important because they point to a strong word class constraint on lexical access: Apparently, only words of the same word class compete for selection.

Table 5
Results of Experiment 2: means and standard errors (*SEM*) for sentence and verbs onsets after related and unrelated noun and verb distractors (all values in ms).

word class	relatedness	mean onset noun (*SEM*)	mean onset (*SEM*)
noun	related	1286 (55)	1086 (53)
noun	unrelated	1256 (55)	1056 (53)
verb	related	1282 (59)	1080 (57)
verb	unrelated	1258 (57)	1056 (55)

Table 6
Estimates of linear mixed effects model on log-transformed sentence onset latencies (Experiment 2). The model included distractor word class, distractor relatedness, their interaction, and trial order (scaled) as fixed effects. Random effects included by-participant slopes and by-item intercepts and slopes for relatedness.

term	estimate	std. error	*t*-value
intercept	6.923	0.034	203.719
relatedness	0.013	0.005	2.804
word class	0.005	0.003	1.415
trial order (scaled)	−0.005	0.000	−41.387
relatedness * word class	0.004	0.003	1.264

Table 7
Estimates of linear mixed effects model on log-transformed noun onset latencies (Experiment 2). The model included distractor word class, distractor relatedness, their interaction, and trial order (scaled) as fixed effects. Random effects included by-participant intercepts and by-item intercepts and slopes for relatedness.

term	estimate	std. error	*t*-value
intercept	7.110	0.030	236.743
relatedness	0.011	0.004	2.612
word class	0.003	0.003	0.999
trial order (scaled)	−0.004	0.000	−41.848
relatedness * word class	0.003	0.003	1.095

This conclusion is consistent with models of sentence generation such as the syntactic traffic cop model, proposed by Dell et al. (2008), where syntactic constraints play a major role in lexical selection and word ordering. Because of the theoretical importance of the findings obtained by Momma and colleagues, we conducted a conceptual replication in a different language (Dutch instead of English). In both of our experiments, we found a main effect of semantic relatedness but no interaction with distractor word class. Thus, in our study semantically related words of a different word class interfered just as much with the

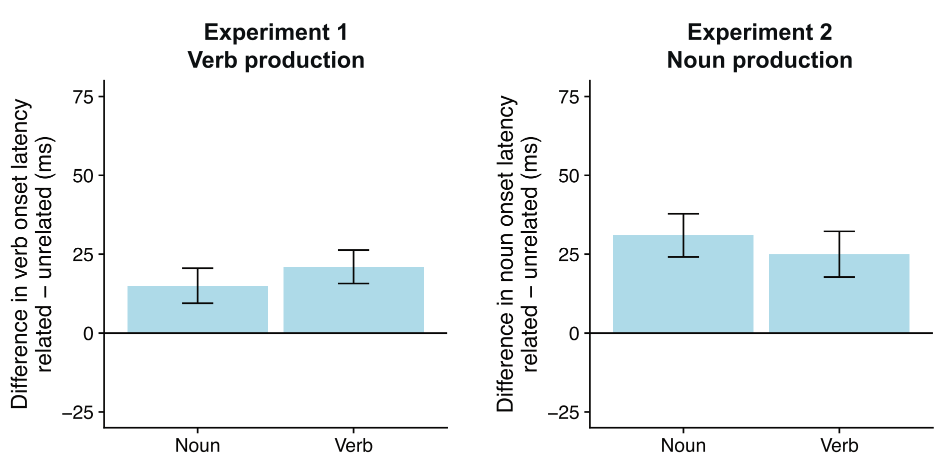

Fig. 2. Semantic interference effects in Experiment 1 (verb production) and Experiment 2 (noun production). The plots depict the differences between the raw (non-transformed) speech onset latencies. Error bars indicate standard errors of the mean.

production of the target nouns or verbs as words of the same word class, contrary to our pre-registered hypothesis. Before turning to cross-linguistic differences that may explain the discrepancy between the two studies, we offer an account of the present results.

4.1. Accounting for the Dutch results

In both experiments, we found a main effect of semantic relatedness between target and distractor. This finding replicates the results Momma and colleagues found in the same-word class conditions of their experiments with the same paradigm. Furthermore, it is consistent with a large body of work using the closely related picture-word interference paradigm (Bürki et al., 2020). The semantic interference effect in picture naming is generally attributed to spreading of activation between related concepts, which affects the activation levels of the associated lexical units and hinders the selection of such units for production. Models differ in where they locate this interference effect: One view, which we discuss first, is that the effect arises during the selection of lemmas, which are grammatically specified word units (Levelt et al., 1999). Another view, also briefly discussed below, allocates the effect at a later stage of processing, namely when a word is selected from a response buffer.

Adopting the lemma selection view, our results can be explained in the following way: When the participants in our Experiment 1 produced an action verb, they identified the action shown in the picture and retrieved the corresponding lexical concept, the verb lemma, and diacritics specifying the surface form. Subsequently, the corresponding morphological and phonological representations were selected followed by the articulatory commands. As the speakers had been asked to memorise the distractor sentence, the lexical concept and lemma of the distractor verb or noun were still active in working memory, and, when semantically related to the target, competed with the target verb for selection. Analogously, when the participants of Experiment 2 produced an action noun, they identified the event in the picture, retrieved a matching lexical concept, the associated noun lemma with diacritics, the corresponding morphological and phonological representations, and the articulatory commands. Again, the presence of a semantically related distractor in working memory created competition at the lemma level and slowed down lexical selection. Note that on this view there are separate lemmas for action verbs and action nouns, whose selection is triggered by distinct lexical concepts. We observed a semantic interference effect across word classes: The selection of action verbs or action nouns was affected to a similar degree by verb and noun distractors. This shows, first, that lexical concepts for action verbs (e.g., *to sing*) activated lexical concepts for closely related action nouns (*the singing*), and, second, that word class did not constrain the competition between lemmas.

In the scenario just described, speakers select lemmas that are specified as verbs or nouns. Alternatively, speakers might initially select lemmas that are not specified for word class and then compose specific lemmas for action nouns and verbs. To elaborate, at the conceptual level there are broad lexical concepts capturing the meaning of words such as *sing* and *whistle*. Each of these lexical concepts is linked to a single lemma. To produce an action verb or an action noun, a speaker selects a lemma and diacritic features that specify the grammatical class, noun or verb, and other relevant features, such as number or tense. These features are then used to construct the morphological representation of the word, which in turn determines the phonological form. A related proposal was put forward by Levelt et al. (1999) for the production of morphologically related homographic word pairs such as “(to) escort”/“(the) escort”. According to their proposal, lemmas are associated with diacritics for grammatical properties such as number or tense. There is also a diacritic for word class, which, in their proposal, is pre-specified as noun or verb. Here we propose that, at least for the Dutch noun/verb pairs we investigated, this feature might not be prespecified but filled in during the grammatical encoding of the words. Since, on this view, the same lemma is initially selected for action nouns and action verbs, there is no word class constraint on the semantic interference effect.

Both of these scenarios presuppose that the semantic interference effect arises during lemma selection (Roelofs, 1992). An alternative is to assign the effect to a later processing step, namely the selection of the appropriate response from a response buffer, which holds both the target and the distractor (Finkbeiner & Caramazza, 2006; Mahon et al., 2007). On this view, a semantically related distractor interferes more with the selection of the target from the response buffer than an unrelated one. The implications of our findings for this view are analogous to the implications discussed above for the lemma selection view: Apparently the match or mismatch in word class between a target and a semantically related distractor does not affect the degree of competition.

In sum, in our experiments with speakers of Dutch we replicated the semantic interference effect, but not the interaction between semantic relatedness and word class congruency. Our findings can be readily explained within mainstream accounts of the semantic interference effect, but, as the interaction was not replicated, they do not support models of utterance planning that posit a strict word class constraint on lexical selection. A plausible source of this discrepancy between the current results and those obtained by Momma and colleagues are cross-linguistic differences between the two languages. We will turn to these after a discussion of methodological differences between the studies.

4.2. Methodological differences between the studies

Although we took care to keep the Dutch translations as close to the English materials as possible, some differences were inevitable. Some of them arose because action nouns are formed differently in the two languages. English action nouns are formed by combining the stem with the affix *-ing,* which is shared with the progressive verbal form. In Dutch, action nouns are formed by combining the stem with the affix *-en*, a form that is shared with the present tense second person plural form. To ensure that the action nouns and verbs were homophones, as in the original study, we changed the pictures to show two agents rather than one and elicited plural forms as target utterances rather than singulars. These necessary changes are clear departures from the original study, but we assume that they did not affect the impact of distractor word class.

Concerning the distractors, the Dutch sentences closely matched the English ones in their structure (see Table 1). However, in Dutch, adverbs are not marked by an affix corresponding to English *-ly*, whereas adjectives in definite noun phrases are marked by the affix *-e.* While the adverb/adjective distinction is clearly marked in Dutch, this distinction may be less salient than in English. Moreover, the possessive pronoun *jullie* (*your*, plural) preceding the action noun in some items is homophonous to the personal pronoun (you). Similarly, in some dialects, possessive *hun* (*their*) can be used instead of *zij* (*they*) as pronoun to mark the third person plural form (*hun zingen* would translate as *they sing*). In these sentences the pronoun was compatible with an upcoming action verb. However, the subsequent adjective disambiguated the phrase as nominal and prevented a verbal interpretation. The speakers recalled the distractor sentences with high accuracy (slightly higher than in the original study), demonstrating that they generated the correct grammatical structures and processed action nouns and verbs as intended. In sum, we consider it unlikely that in our study the word class constraint was absent because the participants did not correctly process the grammatical properties of the distractor sentences.

Finally, there was one other departure from the original study in the materials and target utterances of Experiment 2. In the original study, action nouns were elicited by introducing a colour cue to the action pictures. Participants were briefed to imagine that they were a person with synaesthesia and that thinking of an action simultaneously triggered the perception of a colour. For example, the target sentence for the picture of *whistling* accompanied by the colour cue red was *Her whistling is red*. We simplified the instructions and asked the participants to describe each action as having just started or ended, as cued by a green or red clock, respectively. The resulting target sentences followed the structure *Hun fluiten is begonnen* (*Their whistling has started*). The two participles (*started, ended*) were cued by clearly visible colour cues on the target pictures, as in the original study, to ensure that the processing costs for the verb phrases were minimised and constant across items. However, it is possible that the participles used in the current study encouraged a more verbal interpretation of the target words as compared to the colour attributes: Participles are used to modify verbs, whereas colour adjectives typically modify nouns referring to concrete objects. Thus, through the synaesthesia instruction, Momma and colleagues may have suggested an object-like character of the implied concept and highlighted the word class of the target as being a noun. We used the common participle construction, which did not involve such highlighting. We return to this suggestion below. For now, it is important to recall that the two studies also differed in the results of the two experiments that did have identical designs (Experiment 1 in both studies, where participants produced action verbs). This difference remains to be explained.

4.3. Processing differences between the two languages

If the differences in results between the study by Momma and colleagues and the present study are not due to the specific properties of the materials discussed above, they must be due to more general differences in the way the speakers of American English and Dutch retrieved the target words from their mental lexicon and combined them into the required utterance format. In this section, we consider at which levels of processing such differences might arise.

As explained above, the semantic interference effect originates at the conceptual level, where related concepts activate each other. This mutual activation leads to competition between lemmas during lexical selection or between responses in a response buffer. A straightforward account of the results of both studies is that for speakers of Dutch the lexical concepts for action and object nouns, such as *zij zingen* and *hun zingen*, are tightly linked because they are connected to shared features, whereas for speakers of American English there are no connections between such lexical concepts. While conceptual structures can differ for speakers of different languages (Majid et al., 2004), for the two languages and the specific concepts at issue, fundamentally different conceptual representations seem highly implausible. Yet, there may be subtle differences that impacted the results of two series of experiments. For instance, for some of our experimental items, the conceptual links between action (*singing* as verb) and event (*singing* as noun) concepts might be tighter in one language than in the other.

Alternatively, differences in lemma access might be at play. Above we discussed that speakers might either access lemmas for action verbs and nouns that are specified for grammatical class, or that they might compose lemmas by combining a lemma that is not specified for grammatical class with appropriate diacritics. If both routes to lexical access exist for speakers of Dutch, this should also be the case for speakers of English. However, how likely speakers are to use each route may differ across languages. For example, pre-compiled lemmas for nominalisations may be more readily available for some (or all) items in one language than in the other. Such differences may affect the time course and strength of the activation flow in the system and, with that, the strength of cross-class interference. Consistent with this proposal, there is some evidence that decomposition of compounds may be more likely in Dutch and German than in English (see Zwitserlood (2018) for review and discussion).

Next, we consider the morphological representations of the targets and distractors. The action and object nouns in the two studies were functionally equivalent, but morphologically different: The way in which the forms of action verbs and nouns are generated from the verb stems is different across the two languages. In the English study, verb stems were combined with the suffix *-ing*, which is a marker of the gerund and the progressive form of verbs. In the Dutch study, the verb stems were combined with the suffix *-en,* which marks verb forms in all persons and tenses as plural and is used as the citation form as well. In addition, a homophonous form occurs as the plural marker of many object nouns. For instance, the plural of *hand* (*hand*) is *handen* (*hands*). In other words, the morphological representations of the forms are embedded in different ways in the morphological paradigms of the two languages. They probably also differed in overall frequency and in the function-specific frequencies. For instance, compared to the use as a plural marker of concrete nouns, Dutch *-en* may appear only rarely as a marker of action nouns. English *-ing* may overall be less frequent than Dutch *-en,* and its use as a marker of a noun, rather than verb, may be relatively more common. It is not obvious how such differences by themselves would modulate a semantic interference effect. However, as we will discuss below, they may affect the impact of word class constraints on lexical access.

The final option to consider concerns the word class constraint itself. A straightforward account of the results of both studies is that lexical selection is constrained by word class in English, but not in Dutch. In models of sentence planning, a word class constraint has been proposed to account for an important observation about whole-word speech errors, namely that movement errors typically involve words of the same grammatical class (Fromkin, 1971; Nooteboom, 1973). This constraint on movement errors is taken to result from mechanisms regulating the serial ordering of words in utterances and the availability of abstract sentence representations with slots marked for word class (Garrett, 1975). As explained above, the word class constraint plays a central role in contemporary models of sentence generation, including the syntactic traffic cop model proposed by Dell et al. (2008) and related models by Chang et al. (2006) and Gordon and Dell (2003). Considering the structural similarities across the two languages, it would be very surprising to see that serial ordering was achieved by fundamentally different mechanisms and representations in English and Dutch, and if in English, but not in Dutch, sentence generation were governed by a syntactic traffic cop.

Despite their structural similarities, there are differences in word order flexibility across the two languages (Koster, 1999), which may affect the impact of syntactic constraints during lexical access. For instance, Dutch speakers have more flexibility in the ordering of verb arguments and adjuncts than English speakers, so that the utterances "…dat Jan de man met een bloem begroette" (…that John the man with a flower greeted) and "…dat Jan met een bloem de man begroette" (…that John with a flower the man greeted) are both acceptable. However, utterances with different word orders often differ in information structure. Additionally, the former example contains an attachment ambiguity, which the latter avoids. Consequently, it is not obvious how word order flexibility and the impact of syntactic constraints on utterance planning may be related. On the one hand, it can be argued that a more flexible word order requires looser syntactic constraints because multiple word orders are acceptable. On the other hand, more flexibility in word order may require tighter syntactic constraints, exactly because speakers have to choose from a set of ordering options with slightly different meanings.

Returning to Dell and colleagues' syntactic traffic cop, it is important to keep in mind that the metaphor does not fully capture the underlying theoretical notion and is perhaps a little misleading for two reasons. First, it implies a rigid mechanism that gives or denies classes of words access to a particular sentence position. However, in the model proposed by Dell and colleagues the word class constraint is realised through excitatory and inhibitory connections from syntactic state units to lexical units. For instance, after a determiner has been planned, the weights in the network are changed such that syntactic state nodes boost the activation levels of nouns and reduce the activation levels of verbs, as determiners tend to be followed by nouns rather than verbs. In such an architecture a word class constraint need not be absolute, but can be graded, such that in a given position some word classes are boosted, to different degrees, and others are suppressed. Such an architecture captures the fact that grammatical constraints on word order are often probabilistic. For instance, after a determiner a noun is likely to occur, but so is an adjective or adverb. In other words, the traffic cop prioritises some vehicles over others, without blocking a specific type of traffic altogether.

A second reason why the traffic cop metaphor may be misleading is that it presupposes that all lexical items are unambiguously marked for word class and recognisable as such. For many words, such as concrete nouns and adjectives, this is entirely plausible, but for others the marking may be less clear (e.g., Sasse, 2001). As already indicated, this might in particular be the case for derived forms, such as the action nouns discussed here, which possibly acquire their final grammatical class by combining a stem with appropriate affixes.

To reconcile the results of the study by Momma and colleagues and our own results, we propose that lexical access and sentence planning in English and Dutch follow the same general mechanisms, but that, at least for the word classes relevant here, the impact of syntactic constraints varies. Specifically, semantically related lexical concepts activate each other, regardless of the word class the lexical concepts map onto. Lexical selection is governed by multiple constraints, foremost by the fit of the lexical concept with the intended message, but also by the fit of the lemma with the grammatical structure being built. We assume that the impact of word class constraints varies depending on the strength of the constraint imposed by the current syntactic state (how likely words of a specific word class are to appear in the present context), but also on the clarity of the grammatical markers of the lemmas. In Momma's study of English action nouns and verbs, we observed the impact of a strong word class constraint operating on lemmas clearly marked as nouns or verbs. In our study of Dutch action nouns and verbs, the constraint was not observable (though note that in both experiments the interference effect was numerically slightly stronger for same-class distractor-target pairs than when they belonged to different word classes, consistent with the operation of a weak constraint.) We propose that in our study lemma competition was not limited to a single word class because the classes of action nouns and action verbs were not as clearly marked for grammatical class as the corresponding word classes in the English study. In other words, while the English traffic cop was able to perform their job rigorously and accurately, the Dutch traffic cop struggled to categorise some of the approaching traffic and opted for a lenient policy, admitting everyone. Note that the action nouns and verbs were homophones, and therefore word class errors would not be detected in the participants' overt speech or by their self-monitoring processes. Thus, there was no penalty for lenience.

These considerations are consistent with key assumptions of the model proposed by Dell and colleagues, specifically the assumption of constraints varying continuously (rather than all-or-none) in strength. In addition, we suggest that the relative strength of semantic and syntactic input to lemmas may vary depending on the language and sentence context in which words are produced. Differences in relative strength may be implemented by a learning mechanism such as the *delta rule,* the learning algorithm which in the model by Dell and colleagues adjusts the relative influence of the semantic and syntactic inputs on the lexical unit's activation levels (Dell et al., 2008; Gordon & Dell, 2003). This mechanism could give rise to both language-specific and context-specific differences in the impact of syntactic constraints. Syntactic input may be strengthened (or weakened) depending on the degree of word order flexibility in a language. Furthermore, context-specific variations may arise: Syntactic input may be strengthened in contexts with a highly automatised (or primed) syntactic structure, whereas semantic input may be strengthened in contexts where a word is semantically highly predictable. Assuming that the Dutch nominalised verbs are relatively less frequent than the English equivalents, the syntactic constraint may be weaker in Dutch speakers since there have been fewer occasions to strengthen the input from syntactic state nodes.

A second notion we highlighted in our discussion was that Dutch verbs and nouns were less clearly marked for grammatical class than their English counterparts. As far as we can tell, this is not proposed in the model by Dell et al. Rather, the lexical units in their model are unambiguously categorised as, for instance, determiners, nouns or verbs. In contrast, we propose that a boost from syntactic state nodes to, say, all nouns, is more effective for nouns that are unambiguously marked as such (e.g., *dancer)* than for words that can be nouns or verbs depending on the context (e.g., *dance*). A parameter allowing for different degrees in saliency of word class marking, in combination with the dynamically weighted syntactic and semantic inputs, could account for differences in the impact of syntactic constraints across languages. For example, in a language with flexible word order in which word class marking is opaque, weak input from syntactic state nodes would faintly boost an already-weak word class link, resulting in an altogether loose syntactic constraint. The opposite holds for a language with fixed word order and unambiguous word class marking, where strong syntactic state inputs act on strong word class links, resulting in a tight syntactic constraint.

Evidently, these suggestions need to be tested in further work. It would, in particular, be important to show that within one language, some word classes are more clearly marked than others and that this affects the impact of a word class constraint on lexical selection. A challenge for such work is to vary the word class of distractors while controlling other relevant variables, such as their imaginability and the conceptual relationship to the targets. In future studies one might address this challenge by using large sets of target-distractor pairs and including conceptual and semantic variables as model predictors. It would also be important to assess the proposal that word class constraints are graded, reflecting the strength of word order rules in the language, and that their impact is context dependent. As Momma and colleagues already proposed, in some contexts, speakers may be clearer about the content than the structure of their utterance, which would lead to a weak word class constraint. Furthermore, future studies may use corpus analyses to address the relative frequencies of certain word classes within a given language, as discussed above, and how any cross-linguistic differences might impact the word class constraint. Finally, there may be interindividual variability in sentence planning strategies and in the extent to which syntactic structure is projected. Investigating these issues is essential for moving from the broad claim that the impact of word class constraints is variable to theories that make clear, specific claims about the ways speakers combine lexical and grammatical knowledge when they plan sentences.

5. Conclusions

In two experiments with speakers of American English, Momma and colleagues obtained convincing evidence for a word class constraint in lexical selection. We attempted to replicate their findings in Dutch, a closely related language, using as much as possible translation equivalents of the original materials. We replicated the semantic interference effect but found no evidence for a word class constraint. Given the similarity of the two languages, it seems highly unlikely that lexical selection and sentence generation would follow qualitatively different principles. Instead, we propose that these processes are similar, and, specifically, that lexical selection during sentence generation in both languages is governed by conceptual and grammatical constraints. We suggest that a word class constraint was seen in the English but not the Dutch study because the lemmas of action nouns and verbs were marked more clearly for word class in the former than the latter language.

The results of the two studies highlight how subtle crosslinguistic differences can lead to important processing differences. This implies that psycholinguists need to be cautious when making general claims about language processing based on the results of a single study conducted in a specific language. They should always acknowledge the language under investigation. This is typically done in studies of lesser-studied languages, but often neglected for studies conducted in English, Dutch, or German (van der Burght et al., 2023). On a more general note, we wish to emphasise the importance of performing conceptual replication studies in psycholinguistics, a practice where the field is currently lagging behind compared to other fields within cognitive science (Kobrock & Roettger, 2023).

Open practices statement

Stimulus materials, analysis scripts, and data are available at https://osf.io/gy82u/.

CRediT authorship contribution statement

Constantijn L. van der Burght: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Antje S. Meyer:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

None declared.

Data availability

Data and analysis scripts are freely available on OSF. I have included this in the manuscript (Open Practices Statement).

Acknowledgements

The authors thank Lorien Schipperus and Amelie Ewald for assistance in stimulus preparation and data collection, and Annelies van Wijngaarden for organisational support. They thank Brigitte Bauer for advice on the linguistic properties of the stimuli and Pienie Zwitserlood for valuable discussion. Finally, they would like to thank Geert van der Meulen for adapting the target pictures. The study was funded by the Max Planck Society.

References

Barbiers, S. (2000). The right periphery in SOV languages - English and Dutch. In Svenonius (Ed.), *The derivation of VO and OV* (p. 181). John Benjamins Publishing Co.. https://doi.org/10.1075/la.31.08bar

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Bock, K., & Levelt, W. (1994). Language production: Grammatical enoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984).

Boersma, P., & Weenink, D. (2022). Praat (Version 6.2.14). http://www.praat.org/.

Bürki, A., Elbuy, S., Madec, S., & Vasishth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *Journal of Memory and Language, 114*, 104125. https://doi.org/10.1016/j.jml.2020.104125 (Journal of experimental psychology. Learning, memory, and Cognition 33 3 2007).

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113* (2), 234–272. https://doi.org/10.1037/0033-295x.113.2.234

Costa, A., Alario, F.-X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin & Review, 12*(1), 125–131. https://doi.org/10.3758/bf03196357

Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language & Cognitive Processes, 23*(4), 583–608. https://doi.org/10.1080/01690960801920735

Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you Don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42*(6), 790–796. https://doi.org/10.1016/s0010-9452(08)70419-2

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47* (1), 27. https://doi.org/10.2307/412187

Garrett, M. F. (1975). The analysis of sentence production. *Psychology of Learning and Motivation, 9*, 133–177. https://doi.org/10.1016/s0079-7421(08)60270-4

Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science, 27*(1), 1–40. https://doi.org/10.1207/s15516709cog2701_1

Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. *ArXiv*. https://doi.org/10.48550/arxiv.1802.06893

Iwasaki, N., Vinson, D. P., Vigliocco, G., Watanabe, M., & Arciuli, J. (2008). Naming action in Japanese: Effects of semantic similarity and grammatical class. *Language & Cognitive Processes, 23*(6), 889–930. https://doi.org/10.1080/01690960801916196

Jackendoff, R. (2003). Précis of foundations of language: Brain, meaning, grammar, evolution. *Behavioral and Brain Sciences, 26*(6), 651–665. https://doi.org/10.1017/s0140525x03000153

Janssen, N., Melinger, A., Mahon, B. Z., Finkbeiner, M., & Caramazza, A. (2009). The word class effect in the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology, 63*(6), 1233–1246. https://doi.org/10.1080/17470210903377380

Kobrock, K., & Roettger, T. B. (2023). Assessing the replication landscape in experimental linguistics. *Glossa Psycholinguistics, 2*(1). https://doi.org/10.5070/g6011135

Koster, J. (1999). The word orders of English and Dutch: Collective vs. Individual Checking. *GAGL: Groninger Arbeiten Zur Germanistischen Linguistik, 43*, 1–42.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211–240. https://doi.org/10.1037/0033-295x.104.2.211

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences, 22*(1), 1–38. discussion 38-75.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*(451), 514. https://doi.org/10.3389/fpsyg.2015.01171

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485–495. https://doi.org/10.3758/bf03198265

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical Selection Is Not by Competition: A Reinterpretation of Semantic Interference and Facilitation Effects in the Picture–Word Interference Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences, 8*(3), 108–114. https://doi.org/10.1016/j.tics.2004.01.003

Momma, S., Buffinton, J., Slevc, L. R., & Phillips, C. (2020). Syntactic category constrains lexical competition in speaking. *Cognition, 197*, Article 104183. https://doi.org/10.1016/j.cognition.2020.104183

Nooteboom, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence*.

Pechmann, T., Garrett, M., & Zerbst, D. (2004). The time course of recovery for grammatical category information during lexical processing for syntactic construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30* (3), 723–728. https://doi.org/10.1037/0278-7393.30.3.723

Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 233–243. https://doi.org/10.1037/0278-7393.28.1.233

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Rahman, R. A., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience, 34* (9), 1176–1192. https://doi.org/10.1080/23273798.2019.1599970

Rodríguez-Ferreiro, J., Davies, R., & Cuetos, F. (2014). Semantic domain and grammatical class effects in the picture–word interference paradigm. *Language, Cognition and Neuroscience, 29*(1), 125–135. https://doi.org/10.1080/01690965.2013.788195

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-f

Sasse, H.-J. (2001). Scales between nouniness and verbiness. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals, Volume 1*. De Gruyter Mouton. https://doi.org/10.1515/9783110194036.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*(1), 86–102. https://doi.org/10.1016/0749-596x(90)90011-n

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247–250. https://doi.org/10.1016/j.jml.2004.03.002

van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods, 38*(4), 584–589. https://doi.org/10.3758/bf03193889

van der Burght, C. L., Friederici, A. D., Maran, M., Papitto, G., Pyatigorskaya, E., Schroën, J. A. M., … Zaccarella, E. (2023). Cleaning up the brickyard: How theory and methodology shape experiments in cognitive neuroscience of language. *Journal of Cognitive Neuroscience, 1–22*. https://doi.org/10.1162/jocn_a_02058

van Haeringen, C. B. (1956). *Nederlands tussen Duits en Engels*. Den Haag: Servire.

Vigliocco, G., Vinson, D. P., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition, 94*(3), B91–B100. https://doi.org/10.1016/j.cognition.2004.06.004

Zwitserlood, P. (2018). Processing and Representation of Morphological Complexity in Native Language Comprehension and Production. in *The Construction of Words. Studies in Morphology*, 583–602. https://doi.org/10.1007/978-3-319-74394-3_20