Taylor & Francis Group

The impact of notional number and grammatical gender on number agreement with conjoined noun phrases

Heidi Lorimor, Carrie N. Jackson, Katharina Spalek & Janet G. van Hell

Published online: 25 Apr 2016.

Submit your article to this journal ⤤

View related articles ⤤

View Crossmark data ⤤

Routledge
Taylor & Francis Group

The impact of notional number and grammatical gender on number agreement with conjoined noun phrases

Heidi Lorimor[a], Carrie N. Jackson[b], Katharina Spalek[c] and Janet G. van Hell[d,e]

[a]Department of Languages, Cultures and Linguistics, Bucknell University, Lewisburg, Pennsylvania, USA; [b]Department of Germanic and Slavic Languages and Center for Language Science, Pennsylvania State University, University Park, Pennsylvania, USA; [c]Department of German Language and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany; [d]Department of Psychology and Center for Language Science, Pennsylvania State University, University Park, Pennsylvania, USA; [e]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

ABSTRACT
Morphophonology influences subject–verb agreement in a wide variety of languages. Dominant models of agreement production [e.g. Marking and Morphing, Eberhard, K. M., Cutting, J. C., & Bock, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531–559. doi:10.1037/0033-295X.112.3.531 Competition models, Mirković, J., & MacDonald, M. C. (2013). When singular and plural are both grammatical: Semantic and morphophonological effects in agreement. *Journal of Memory and Language, 69*, 277–298. doi:10.1016/j.jml.2013.05.001 posit explanations for morphophonological effects that depend on ambiguity. The present study uses sentence completion tasks in Dutch (Experiment 1) and German (Experiment 2) that manipulate notional number and grammatical gender with conjoined noun phrases to investigate how morphophonology affects number agreement. Results show that speakers of both languages produced more singular agreement with items construed as more notionally singular, and with items containing two nouns with the same grammatical gender, even though, prima facie, grammatical gender should be irrelevant for subject–verb number agreement in these languages. Experiment 2 showed that the grammatical gender effect was not driven by morphophonological ambiguity. These results provide novel insight into how morphophonology, via cue-based retrieval, can affect subject–verb number agreement.

A driving question in psycholinguistic research is the degree to which language processes are influenced by "irrelevant" information. For instance, subject–verb number agreement, which is often viewed as depending solely on the grammatical number of the subject noun (Quirk, Greenbaum, Leech, & Svartvik, 1985), is actually sensitive to semantic information (e.g. Eberhard, 1997; Morgan, 1984; Pollard & Sag, 1994; Vigliocco & Hartsuiker, 2002). Other factors, like morphophonology, also influence subject–verb agreement (e.g. Foote & Bock, 2012; Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008).

The factors that influence subject–verb agreement have been primarily studied through the phenomenon of agreement "attraction" (Bock & Miller, 1991), which refers to the increase in subject–verb agreement errors in the presence of non-subject, attractor nouns. Agreement attraction occurs in number agreement, as in "the label on the bottles are …" (Vigliocco, Hartsuiker,

Jarema, & Kolk, 1996). Attraction also occurs with gender agreement, as in "*la-F.SG vista-F.SG de los-M.PL puertos-M.PL es-SG bonito-M.SG* 'the view of the ports is pretty'" (Antón-Méndez, Nicol, & Garrett, 2002), as the adjective *bonito* agrees in gender with the masculine attractor noun, rather than with the feminine subject head noun.

Two dominant models for subject–verb agreement are the Marking and Morphing model (Eberhard, Cutting, & Bock, 2005), and the Competition model (e.g. Mirković & MacDonald, 2013), which can account for a wide range of effects, including agreement attraction and the impact of notional number. Both models also include mechanisms through which morphophonology influences subject–verb agreement. To date, the effects of morphophonology on agreement have been somewhat elusive because, apart from a few constructions, like those involving collectives (Haskell & MacDonald, 2003) or quantifiers

(Mirković & MacDonald, 2013), there is little variability in agreement patterns, leading to "floor effects" (Haskell, Thornton, & MacDonald, 2010). One construction that does exhibit variable agreement is the conjoined noun phrase (NP) (e.g. *the salt and pepper*), especially when both of the conjoined nouns are singular (Lorimor, 2007). The current study uses conjoined NPs with a range of notional number valuations to investigate the role of morphophonology – manipulated here via gender morphology on determiners in the NP – on subject–verb agreement in Dutch (Experiment 1) and German (Experiment 2). By manipulating notional and morphophonogical information with a construction that is likely to show a wide range of agreement patterns, the present experiments enhance our understanding of the mechanisms through which morphophonology affects subject–verb agreement (e.g. Antón-Méndez & Hartsuiker, 2010; Badecker & Kuminiak, 2007; Foote & Bock, 2012; Lago, Shalom, Sigman, Lau, & Phillips, 2015; Mirković & MacDonald, 2013), and address the degree to which current models of agreement can account for such effects.

Conjoined NPs

Agreement with conjoined NPs is a particularly complex question, both because conjoined NPs tend to be more flexible in their agreement behaviour than many other constructions, and because languages differ in how agreement works with conjoined NPs (Badecker, 2007; Corbett, 2006). In English, speakers frequently use singular verbs with conjoined NPs (Lorimor, 2007), and singular agreement is also frequently used with conjoined NPs in other languages (Corbett, 2006). In addition, some languages allow for "closest-conjunct agreement", which is agreement with the conjunct that is closest to the verb, rather than with the conjoined NP as a whole. However, other languages, including English, Dutch and German, do not allow closest-conjunct agreement as a grammatical option (van Koppen, 2005), so if singular verb agreement arises in these languages that do not allow for closest-conjunct agreement, this can provide important insight into the underlying mechanisms that compute agreement for the whole subject NP during language production.

Morphophonological effects on agreement

Investigating the impact of morphophonology on agreement provides valuable insights into the underlying mechanisms that drive agreement production. To date, morphophonological effects on agreement have been primarily observed in languages that exhibit some

degree of morphophonological ambiguity. For instance, the Dutch determiner *de* is ambiguous for number, as it is used for both singular common-gender nouns and all plural nouns in Dutch. In contrast, *het* is morphologically unambiguous and is only used with singular neuter nouns. Previous experiments have shown that speakers produce more agreement attraction errors when the determiner associated with the subject head noun is ambiguous for number (*de*) than when the subject head noun is unambiguously singular (*het*) (Antón-Méndez & Hartsuiker, 2010; Hartsuiker, Schriefers, Bock, & Kikstra, 2003).

Ambiguity can also explain morphophonological effects from homophony within a case system. For instance, in Serbian, the feminine genitive plural is homophonous with the nominative singular ending; Mirković and MacDonald (2013) found that with quantifier expressions, the presence of a genitive plural noun that was homophonous with a nominative singular form was associated with increased rates of singular verb agreement (see also Hartsuiker et al., 2003, for similar effects of case marking in German and Badecker & Kuminiak, 2007, for examples with gender attraction in Slovak). An analysis of Russian also showed an increase in number agreement attraction with non-subject plural nouns that were homophonous with nominative plural forms (Lorimor, Bock, Zalkind, Sheyman, & Beard, 2008).

Morphophonological ambiguity can vary across language varieties. Foote and Bock (2012) compared the rate of agreement attraction in languages that are less morphologically complex (Dominican Spanish and English) to a language variety that has more robust number morphology (Mexican Spanish). They found that speakers of Dominican Spanish and English were more likely to make attraction errors than speakers of Mexican Spanish. Foote and Bock interpreted their data within the Marking and Morphing model, arguing that the unambiguous number morphology in Mexican Spanish led to more grammatical agreement overall, while the sparser morphophonology in Dominican Spanish led to a higher rate of agreement attraction.

Ambiguity has therefore been incorporated into how current models of agreement account for morphophonological effects. Within Marking and Morphing, the role of morphophonology comes primarily through the sparseness or ambiguity of a feature set (Antón-Méndez & Hartsuiker, 2010). For example, in Dutch, each instance of the determiner *het* within a subject NP would contain an additional singular specification, which would lead to more singular agreement than when the subject NP contains a number-ambiguous *de*-determiner (Antón-Méndez & Hartsuiker, 2010). In a

Competition model, these effects are explained through distributional patterns (Mirković & MacDonald, 2013). For Dutch, the Competition account can be modelled as follows: When speakers encounter the *het* determiner, this is frequently in the context of a singular verb. However, when they encounter the *de* determiner, they are likely to encounter either a singular or plural verb. A speaker's prior experience with these distributional patterns influences the relative activation and subsequent selection of competing word forms (e.g. *is/are* …). Therefore, both of these models, which use ambiguity to explain morphophonological effects, would predict less singular agreement when determiners are ambiguous for number than when the determiners are unambiguously singular.

Cue-based retrieval

In addition to these afore-mentioned models of agreement (Competition [Mirković & MacDonald, 2013] and Marking and Morphing [Eberhard et al., 2005]), other work has focused on the role of cue-based retrieval in agreement production and comprehension (Lago et al., 2015; Tanner, Nicol, & Brehm, 2014; Wagers, Lau, & Phillips, 2009; see also Badecker & Kuminiak, 2007). In a follow-up study on Foote and Bock's (2012) data from Mexican and Dominican Spanish, Lorimor, Jackson, and Foote (2015) showed that, while the robustness of the number morphology in Mexican versus Dominican Spanish could account for a significant difference in the rate of agreement attraction between the two languages, grammatical gender played an additional role in number agreement. Specifically, when speakers of both varieties of Spanish produced overt subjects (Experiment 1; Foote & Bock, 2012), they made fewer errors in number agreement when the head and local nouns had different grammatical genders than when they had the same gender. Given that the reduction in number agreement errors occurred in both language varieties when the head and local nouns had different grammatical genders, this demonstrates that the "sparseness" explanation is unable to account for the full pattern of morphophonological effects in Spanish. Instead, Lorimor et al. (2015) argue that grammatical gender can serve as a cue that facilitates correct retrieval of the subject head noun when the speaker is planning the verb. They propose that the agreement patterns in Foote and Bock (2012) are best understood by combining a model like Marking and Morphing (Eberhard et al., 2005), which can explain how the form of the verb is originally computed, with a cue-based retrieval model, in which a speaker would use the gender cues

from the head and local nouns to verify that the subject and verb agree.

For conjoined NPs, the effects of cue-based retrieval have yet to be explored. On the one hand, cue-based retrieval might predict some degree of agreement with the closest NP, as the noun nearest to the verb serves as a plausible agreement controller (Thornton & MacDonald, 2003). If this is the case, we might see more singular verb agreement when the closest NP is an unambiguously singular *het*-noun, as the morphophonological cue on the nearest conjunct would lead to more singular agreement, compared to when the closest NP has a number-ambiguous determiner, *de*. However, Lago et al. (2015) found that the robustness of number morphology of a local noun did not affect agreement attraction in comprehension, as attraction effects were similar in English and Spanish, even though Spanish has more robust morphology. Further, using reading-time measures, they showed that only a small portion of readers slowed down at the verb, which they interpreted as evidence that cue-based retrieval is an error-driven mechanism that only occurs when a listener encounters an ungrammatical verb.

On the other hand, even if there is no effect of the morphophonological ambiguity on the local noun, cue-based retrieval might predict that, by having determiners that reflect two different genders within the conjoined NP, this would lead to more plural agreement, as the two distinct determiners would remind the speaker that there are two distinct nouns in the subject NP. This is consistent with work that shows how grammatical gender can help speakers locate the correct subject nouns, both for production (Lorimor et al., 2015) and for comprehension (Adani, van der Lely, Forgiarini, & Guasti, 2010; Belletti, Friedmann, Brunato, & Rizzi, 2012). The mechanism of cue-based retrieval would not replace other models of agreement, such as the Competition model (Mirković & MacDonald, 2013), or Marking and Morphing (Eberhard et al., 2005), but would operate as an error-driven mechanism once the form of the verb had been computed. For example, a singular verb form might be generated by the Marking and Morphing model because the NP is notionally singular and because conjoined NPs often occur with singular agreement (Lorimor, 2007). However, as the speaker is planning to produce the verb form, they may verify that they are indeed producing the correct form of the verb by consulting the grammatical features of the subject NP that the speaker is holding in content-addressable memory (Lewis & Vasishth, 2005). If the speaker is holding two distinct gender-marked determiners in content-addressable memory, then we would predict less singular agreement with conjoined NPs when the

conjoined nouns have determiners expressing different grammatical gender, compared to when the conjoined nouns have the same grammatical gender.

Notional effects on agreement with conjoined noun phrases

To study how morphophonology affects agreement, we will tap into the variability in agreement that can be achieved through the manipulation of notional number. In complex NPs, notional number is often operationalised through distributivity (e.g. Eberhard, 1999; Humphreys & Bock, 2005). With conjoined NPs, notional number is operationalised through degrees of "coalescence" (Lorimor, 2007), or notional linkage. Essentially, coalescence means that speakers can think about conjoined NPs as together forming a larger unit. For a phrase like *the wind and the rain*, the speaker can think about the wind and rain as one larger "weather-based" unit. Other phrases, like *the sun and the moon,* are less likely to coalesce, as these objects are more likely to be viewed independently of each other.

Lorimor (2007) found that, in English, conjoined NPs exhibit a range of agreement patterns and that notional information influences speakers' decisions to use singular or plural verbs. In an oral sentence completion task, rates of singular agreement ranged from 19% (for count nouns, e.g. *the violin and viola*) to 67% (for deverbal nouns, e.g. *the preparation and cooking*) when both of the conjoined nouns were singular. Given that Dutch and German are more morphologically complex than English, we might expect lower rates of singular agreement in Dutch and German than in English (cf., Foote & Bock, 2012; Lorimor et al., 2008), but even a reduced notional effect may provide enough variability to investigate the role of gender morphophonology in number agreement.

Present study

In the present experiments, we used conjoined NPs in Dutch (Experiment 1) and German (Experiment 2) as a window into how morphophonology affects agreement in the presence of notional differences. All of the nouns within the conjoined NPs were grammatically singular to maximise the likelihood of eliciting singular verb agreement. The variation in notional number was achieved by including conjoined NPs of different semantic types: from animate and count nouns, which are more notionally plural, to mass and deverbal nouns, which are more notionally singular. In line with previous findings from English (Lorimor, 2007), we would predict more singular agreement for conjoined NPs involving mass

and deverbal nouns than with conjoined NPs involving animate and count nouns.

The morphophonological manipulation was achieved through varying the gender of nouns with their corresponding gender-marked determiners in conjoined NPs. In Dutch, this involved using common-gender nouns with the determiner *de*, which is ambiguous for number, and neuter-gender nouns with the determiner *het,* which is unambiguously singular, in all four possible combinations, *de-de, de-het, het-de* and *het-het*. In German, there are three gender-marked determiners. Two are unambiguously singular: the masculine determiner (*der),* and the neuter determiner (*das*). The feminine determiner (*die*) is ambiguous for number. We can manipulate gender match and the presence of morphophonological ambiguity by including items in all gender conditions. This was accomplished by creating two variables related to grammatical gender. The first variable, "gender match" separated all of the conjoined NPs with matching grammatical gender (in Dutch, *het-het, de-de*; in German *das-das, die-die, der-der*) from those with mismatching gender. The second variable, "closest NP" coded the number ambiguity of the closest (2nd) NP. (In Dutch, *het-het* and *de-het* have the unambiguous *het*-NP in the 2nd position; in German this would consist of any item with the unambiguous *das*-or *der*-NP in the 2nd position.) If number ambiguity alone drives morphophonological effects on agreement, then we would expect the most singular agreement when the determiners are unambiguously singular (*het-het* (in Dutch) and *der-das, das-der, das-das, der-der* (in German)) and the least singular agreement for items that contain number-ambiguous determiners (*de-de* (in Dutch) and *die-die* (in German)). However, if the grammatical gender of the individual nouns within the conjoined NP exhibits other effects on number agreement, this could provide further information about how cue-based retrieval works in language production. On the one hand, it is possible that having an unambiguously singular determiner on the closest NP would lead to more singular agreement, suggesting an overall strategy for speakers to agree with the closest noun available. On the other hand, if there is more plural agreement whenever the conjoined NPs are mismatched for gender, independent of whether the determiners are ambiguous or unambiguously marked for number, this could be explained by an error-driven cue-based retrieval mechanism. This is because the two distinct gender features would remind the speaker to retrieve both nouns in the conjoined NP.

Corpus-based pre-study

To determine whether Dutch has sufficient variability in subject–verb agreement to investigate morphophonological

effects, and to have a point of comparison for our experimental results, we conducted a corpus-based analysis with conjoined NPs. The corpus was a subset of the *Corpus Gesproken Nederlands* "The Corpus of Spoken Dutch" that has been tagged for part of speech (Hoekstra et al., 2003). We limited the analysis to conjoined NPs in which both nouns were either a simple count or animate noun, or both nouns were either a simple mass or deverbal noun, to approximate the stimuli in the sentence completion experiment. This gave us a total of 82 tokens, of which 18 were used with singular verbs (22.0%). We then separated the 82 items by notionality and found that 5 of 53 animate/count conjoined NPs (9.4%) and 13 of 29 mass/ deverbal conjoined NPs (44.8%) were used with singular verbs. This shows that singular agreement with conjoined NPs occurs in spontaneous Dutch speech and that the rate of singular agreement is higher for mass/deverbal conjoined NPs than count/animate conjoined NPs, ($p < .001$, Fisher's Exact Test), consistent with notional effects.

Experiment 1: Dutch sentence completion task

In Experiment 1, participants completed an oral sentence completion task using Dutch conjoined NPs to investigate whether grammatical gender has an impact on number agreement in Dutch, and if so, whether this effect is driven by number ambiguity on the determiner. We then compared the results of our study to a computational implementation of the Marking and Morphing model (Eberhard et al., 2005), to determine the degree to which Marking and Morphing, with its reliance on morphophonological ambiguity, can account for the observed effects of morphophonology in Experiment 1.

Method

Participants

Fifty-five Dutch native speakers in the Netherlands completed the experiment. Due to technical difficulties, responses from one participant were not recorded. All results are based on data from the remaining 54 participants (44 female; 10 male). The mean age of participants was 19.3 years (SD = 1.7; range: 18–24).

Materials

All experimental items consisted of a conjoined NP containing two singular nouns along with their definite article, joined with the conjunction *en* "and". Items varied according to grammatical gender match, with conjoined NPs containing either two nouns that matched in grammatical gender (*de-de* or *het-het*) or two nouns with different gender (*het-de* or *de-het*).

Items also varied according to the notionality of the conjoined NP as a whole (less notionally singular vs. more notionally singular). Based on Lorimor (2007), less notionally singular conjoined NPs consisted of two animate or two simple count nouns. More notionally singular conjoined NPs consisted of two mass or two deverbal nouns. See Table 1 for sample items.

The final set of 80 experimental NPs was chosen from a pool of 96 candidate NPs based on the results of three norming tasks: a notionality rating task, a sensibility rating task and an article-use rating task.

In the notionality rating task, 103 Dutch native speakers – none of whom participated in the sentence completion task – rated whether they were more likely to replace each phrase with the singular personal pronoun *het* "it" or the plural personal pronoun *ze* "they" on a 5-point scale (1 = *alleen het* "only it"; 3 = *allebei* "both"; 5 = *alleen ze* "only they"). The 96 candidate experimental items were divided into four lists so that each participant rated 24 experimental items and 8 filler items, which consisted of conjoined NPs containing both plural and singular nouns, presented in randomised order.

In the sensibility rating task, 100 participants from the notionality rating task rated how sensible the conjoined NPs sounded on a 5-point scale (1 = *onzinnig* "nonsensical"; 5 = *zinnig* "sensible"). The experimental items were divided into four lists so that participants rated 24 experimental items and 8 filler items, which consisted of conjoined NPs containing plural and singular nouns, that were either sensible (e.g. *De diamanten en het platina* "the diamonds and the platinum") or nonsensical (e.g. *De rokken en de munten* "the skirts and the coins"), presented in randomised order. Lists were distributed so that participants did not rate the same set of experimental items in both the notionality and sensibility rating tasks.

In the article-use rating task, 20 new Dutch native speakers rated how natural each candidate item sounded with both definite articles on a 6-point scale (1 = *zeer onnatuurlijk* "very unnatural"; 6 = *zeer natuurlijk* "very natural"), which were presented in randomised order. We included this task because in Dutch, when both nouns in a conjoined NP have the same grammatical gender, it is acceptable to eliminate the second article (e.g. *de-COM thee-COM.SG en koffie-COM.SG* "the tea and coffee") whereas this is not possible when the nouns have different gender (e.g. **het-NEUT.SG brood-NEUT.SG en boter-COM.SG* "the bread and butter"). We wanted to ensure that both same and different-gender conjoined NPs sounded equally natural when they included definite articles with both nouns.

For the critical trials in the experiment, 80 conjoined NPs were chosen so that there were 40 animate or

Table 1. Sample stimuli for Dutch sentence completion task (Experiment 1).

Gender condition		Animate/count nouns (less notionally singular)	Mass/deverbal nouns (more notionally singular)
Different gender	Het-de	Het meer en de rivier	Het meel en de suiker
		"The lake and the river"	"The flour and the sugar"
	De-het	De vork en het mes	De salami en het spek
		"The fork and the knife"	"The salami and the bacon"
Same gender	De-de	De tafel en de stoel	De ketchup en de mayonaise
		"The table and the chair"	"The ketchup and the mayonnaise"
	Het-het	Het hoofd en het hart	Het koper en het ijzer
		"The head and the heart"	"The copper and the iron"

count items and 40 mass or deverbal items. Within each notional category, 20 items had different grammatical gender (animate/count: 10 *het-de* and 10 *de-het* items; mass/deverbal: 12 *het-de* and 8 *de-het* items), and 20 items had the same grammatical gender (animate/count: 9 *de-de* and 11 *het-het* items; mass/deverbal: 13 *de-de* and 7 *het-het* items). See Table 2 for the descriptive results of the three norming tasks. ANOVAs treating notionality (animate/count vs. mass/deverbal) and gender match (same vs. different gender) as between-item variables revealed no significant effects or interactions for the sensibility or article-use ratings (all *p*s > .05). There was a main effect of notionality in the notional rating task because animate and count items were rated as less notionally singular than mass and deverbal items ($F(1, 76) = 75.34$, $p < .001$), but there were no other significant effects or interactions (ps > .6).

The 80 conjoined NPs were split into two experimental lists, such that participants saw 10 animate/count items with the same gender, 10 animate/count items with different genders, 10 mass/deverbal items with the same gender and 10 mass/deverbal items with different genders. These experimental NPs were presented in a randomised order along with 112 filler NPs. Filler NPs included conjoined NPs that contained at least one plural noun (e.g. *Het debat en de vraagstukken* "The debate and the statements"), complex NPs (e.g. *Het pad in de bergen* "the trail in the mountains") and simple singular and simple plural NPs (e.g. *Het groene gordijn* "the green curtain"). Across all items, an equal number of NPs were – prescriptively speaking – grammatically plural and grammatically singular. All experimental and filler NPs were recorded by a female Dutch native speaker.

Procedure

Participants were tested individually in a quiet room, and their responses were digitally recorded. Participants received both aural and written instructions. They were told to listen to each phrase (preamble), to repeat each phrase exactly as they heard it and to then complete the sentence by stating where the things are (Lorimor, 2007). Participants were instructed to use the copula verb in either the present or past tense (i.e. *is/zijn* "is/are"; *was/waren* "was/were") and to respond as quickly and fluently as possible. In the instructions, participants were given several example completions (e.g. *in de stad* "in the city", *op de maan* "on the moon") and one complete example. Participants received feedback if they repeatedly used a verb besides the copula or if their response did not specify a location, but they received no feedback on their repetition of the preamble or the use of singular or plural verbs in their completions. The experiment began with 10 practice items.

The preambles were presented using E-Prime version 2.0 (Schneider, Eschman, & Zuccolotto, 2002). Participants first saw a fixation point for 500 ms and then heard the recorded preamble. Then an exclamation point appeared on the screen, prompting participants to repeat the preamble and complete the sentence. The experimenter advanced between trials manually with a mouse click. Sessions lasted 15–20 min.

Scoring

Responses from all participants were transcribed and coded as singular, plural or miscellaneous. A second coder scored 5% of the data, with an interrater reliability of 97.6%. Miscellaneous responses, which were scored

Table 2. Mean ratings (standard deviation) and range for notionality, sensibility and article use for Dutch sentence completion task (Experiment 1).

	Notionality		Sensibility		Article use	
	M (SD)	range	*M* (SD)	range	*M* (SD)	range
Animate/count nouns						
Same gender	3.6 (0.6)	2.7–4.6	3.7 (0.3)	3.2–4.3	4.6 (0.3)	4.1–5.2
Different gender	3.7 (0.5)	3.0–4.6	3.7 (0.5)	3.0–4.6	4.8 (0.2)	4.3–5.2
Mass/deverbal nouns						
Same gender	2.7 (0.4)	2.0–3.4	3.7 (0.3)	3.1–4.3	4.6 (0.3)	3.7–5.1
Different gender	2.7 (0.3)	2.1–3.4	4.0 (0.4)	3.2–4.5	4.7 (0.3)	3.9–5.1

Table 3. Distribution of response type by scoring category and percentage of plural versus singular verbs for non-miscellaneous responses in Dutch sentence completion task (Experiment 1).

Response type			
	Plural (% plural)	Singular (% singular)	Miscellaneous (misc. plural, misc. singular)
Animate/count nouns			
Het-de	219 (97%)	7 (3%)	44 (31, 3)
De-het	226 (98%)	4 (2%)	40 (32, 0)
De-de	210 (98%)	4 (2%)	30 (25, 0)
Het-het	243 (98%)	5 (2%)	48 (38, 3)
Mass/deverbal nouns			
Het-de	221 (90%)	24 (10%)	79 (54, 5)
De-het	152 (85%)	27 (15%)	37 (26, 3)
De-de	240 (85%)	42 (15%)	68 (50, 5)
Het-het	106 (68%)	51 (32%)	33 (15, 11)

separately, included instances where the participants did not correctly repeat the preamble, hesitated or repeated part of the preamble, inserted a filled pause, responded with a verb other than the copula, or failed to complete the sentence. There were 1617 (74.9%) plural responses (898 with animate/count items; 719 with mass/deverbal items), 164 (7.6%) singular responses (20 with animate/count items; 144 with mass/deverbal items), and 379 (17.5%) miscellaneous errors (162 with animate/count items; 217 with mass/deverbal items). All responses are presented in Table 3.

Data analysis

Analyses were conducted using mixed-effect logistic regression models (Jaeger, 2008) with the lme4 package in R version 2.15.3 (R Development Core Team, 2013). Although we included both animate/count and mass/deverbal items to capture the full range of notionality with conjoined NPs (cf. Lorimor, 2007), as seen in Table 3, agreement with conjoined NPs containing animate/count items was almost uniformly plural. Therefore, the primary statistical comparison reported here included only the mass/deverbal items (see Foote & Bock, 2012, for a similar procedure; see online supplemental material for the results of a preliminary analysis directly comparing animate/count items to mass/deverbal items). Gender match (different vs. same gender) was entered as a fixed effect, contrast coded with the values −.5 and .5. As notionality is not just a categorical distinction, but rather exists on a continuum, even within mass/deverbal conjoined NPs, we entered each item's rating from the notionality rating task as a fixed effect, centred at the sample mean of 2.74. In order to fully capture the effect of number ambiguity, we also included the ambiguity of the determiner on the closest NP (ambiguous-*de* vs. unambiguous-*het*) as a fixed effect, contrast coded with the values −.5 and .5. Finally, to control for any differences in sensibility

or article-use ratings between individual items, we entered these as continuous variables, centred at the sample means of 3.82 and 4.64, respectively.

Results

In an initial model with the mass/deverbal items, we entered the interaction terms notionality × gender match, notionality × closest NP and gender match × closest NP, to investigate whether the magnitude of any effect varied according to another primary variable of interest. However, none of these interactions improved the overall model fit, as measured by a Chi-square goodness-of-fit test comparing the −2 restricted log likelihood values of the model including the variable to a version of the model without the variable (Quené & van den Bergh, 2008; notionality × gender match $p = .158$; all other $ps > .5$), so they were not included in the final model reported here (see online supplemental material for the model with interaction effects). The random effect structure included random intercepts for items and participants, and by-participant random slopes for the closest NP. The inclusion of additional random slopes did not significantly improve the model fit (all $ps > .6$), however all results reported here also hold if the model includes the maximal uncorrelated random effect structure supported by the design.

As seen in Table 4, there were no significant effects of sensibility or article-use ratings. There was a significant effect of notionality because the proportion of singular responses increased on items rated as more notionally singular (i.e. items with lower notionality rating scores). There was a significant effect of gender match because the proportion of singular responses was higher for items with the same gender (i.e. *de-de* or *het-het*) than items with different gender (i.e. *de-het* or *het-de*). There was also a marginally significant effect of closest NP because the proportion of singular responses was

Table 4. Summary of the mixed logit model for Dutch mass/deverbal conjoined NPs (Experiment 1).

Predictor	Parameter estimates		Wald's test of individual coefficients		Δ(-2Λ)-test	
Fixed effects	Estimate	Std. error	z-value	Pr (>|z|)	χ²	p
(Intercept)	−2.01	0.20	−10.11	<.001		
Sensibility	0.50	0.44	1.14	.256	1.26	.261
Article use	0.03	0.45	0.08	.940	0.01	.941
Notionality	−0.89	0.40	−2.20	.028	4.65	.031
Gender match	0.49	0.15	3.28	.001	9.67	.002
Closest NP	0.31	0.16	1.90	.058	3.34	.068

higher when the determiner on the second noun was *het* than when the determiner on the second noun was *de*.

Marking and Morphing model implementation

Experiment 1 showed an effect of gender match, with more singular agreement for same-gender conjoined NPs than for different-gender conjoined NPs. At first glance, it is not clear how this effect could be explained by Marking and Morphing, which accounts for morpho-phonological effects via ambiguity. However, Marking and Morphing has a computational implementation (Eberhard et al., 2005), which calculates the predicted proportion of plural verbs for each item, allowing us to directly compare the predictions of the model to our experimental results. In this way, we can test the model's predictions for how grammatical gender should affect number agreement in Dutch, and make predictions that are specific to our set of experimental items. For ease of exposition, and to parallel the presentation of our experimental and corpus results, we will convert the results of the model from proportion plural to proportion singular agreement.

In the Marking and Morphing model, each component of the NP (determiners, subject head nouns and non-head (local) nouns) has a singular, plural, or unmarked specification. Grammatical number specifications range from −1 (for items marked as singular) to +1 (unambiguously plural). These values are obtained from corpus analyses (see Eberhard et al., 2005, for an explanation), based on how often each element co-occurs with singular or plural verbs. Not all elements are equally important for determining agreement, so the "weight" of each element is included as a multiplier. For example, head nouns have a weight of 18.31, while determiners have a weight of .28. Notional number valuations of each NP are incorporated into the model via an additional term that is summed along with the values derived from the components of the NP.

In our implementation, the grammatical number specification for each conjoined NP was calculated separately, using the same equation as Antón-Méndez and

Hartsuiker (2010) (see also Equations 9 and 10 in Eberhard et al., 2005):

$$S(r) = S(n) + w_{NP1}(S(m)_{N1} + (w_{Det1} * S(m)_{Det1}))$$
$$+ w_{en} * S(m)_{en} + w_{NP2}(S(m)_{N2} + (w_{Det2} * S(m)_{Det2}))$$

$S(r)$ represents the total specification of plural for the entire conjoined NP, with higher $S(r)$ values indicating a greater probability of producing a plural verb. The proportion of plural agreement predicted by the model is calculated using a logistic function with a bias parameter $b = -3.42$ (Equation 5 in Eberhard et al., 2005).

$$P_{plur} = \frac{1}{1 + \exp^{-[S_r + b]}}$$

$S(n)$ represents the notional number value for the conjoined NP, which was operationalised as the z-scores of the ratings from the notional rating task described above (see Mirković & MacDonald, 2013, for a similar procedure). Higher $S(n)$ values represent greater notional plurality. Following the procedures outlined in Eberhard et al. (2005), the $S(m)$ values for N1 and N2 were set at 0 or −1, representing the grammatical number value for each individual noun in the conjoined NP. Most of the nouns (singular count, mass and deverbal) had values of 0, as they are unmarked for number. Invariant singular nouns had values of −1, as they are unambiguously singular by virtue of having no contrasting plural form.[1] For the nouns within the conjoined NP, the weights w_{NP1} and w_{NP2} were set at 1.39, the weight established by Eberhard et al. (2005) for NPs that are embedded within the subject NP (see also Antón-Méndez & Hartsuiker, 2010).

The $S(m)$ values for the determiners were set at 0 for the determiner *de*, because it is ambiguous for number, and −1 for the determiner *het*, because it is unambiguously singular (following Antón-Méndez & Hartsuiker, 2010). The weights w_{Det1} and w_{Det2} were set at 0.28, following the weights established by Eberhard et al. (2005) and used by Antón-Méndez and Hartsuiker (2010) for determiners.

We include a number specification and weight for *en* "and" to capture the agreement properties of the conjoined NP, as neither of the (singular) nouns within the

Table 5. Model parameters for Marking and Morphing model implementation, with proportion of singular agreement (Model) compared with data from Experiment 1.

	$w_{NP1/NP2}$	$S(m)_{N1/N2}$	$w_{Det1/Det2}$	$S(m)_{Det1}$	$S(m)_{Det2}$	Model	Exp. 1
Animate/count nouns							
Het-de	1.39	−1 or 0	0.28	−1	0	.011	.031
De-het	1.39	−1 or 0	0.28	0	−1	.011	.017
De-de	1.39	−1 or 0	0.28	0	0	.008	.019
Het-het	1.39	−1 or 0	0.28	−1	−1	.018	.020
Mass/deverbal nouns							
Het-de	1.39	−1 or 0	0.28	−1	0	.104	.098
De-het	1.39	−1 or 0	0.28	0	−1	.081	.151
De-de	1.39	−1 or 0	0.28	0	0	.055	.149
Het-het	1.39	−1 or 0	0.28	−1	−1	.294	.325

$S(n)$: −0.39–0.44 w_{en}: 10.42 $S(m)_{en}$: 0.78

conjoined NP functions as a traditional "head-noun". In treating the conjunction *en* "and" as the head of the conjoined NP, we do not intend to take a particular theoretical stance on the internal syntactic structure of conjoined NPs. Rather, we do so simply to capture notional and grammatical properties of the conjoined NP as a whole in our model implementation. We based $S(m)_{en}$ on the proportion of plural agreement in our corpus data (.78) (cf., Mirković & MacDonald, 2013), from the subset of 82 conjoined NPs with two simple count/animate or two simple mass/deverbal nouns. The weight w_{en} was set at 10.42. This value was calculated by a stepwise function that minimised the root mean squared error (RMSE) of the model, which is similar to how the weight for each element was originally set in Eberhard et al. (2005).

Model parameters are summarised in Table 5. Table 5 also compares the predicted proportion of singular agreement to the experimental production data. This model implementation predicts between 1% and 29% singular agreement across all notional and gender manipulations. The correlation between the predicted and observed proportion of plural verbs is 0.60, and the prediction error (RMSE) is 0.099. This contrasts with RMSE values ranging from 0.032 to 0.077 in Eberhard et al. (2005) (but see Mirković & MacDonald, 2013, for higher RMSE values ranging from 0.40 to 0.69).

Inspection of the predicted results in Table 5 reveals that, for mass/deverbal conjoined NPs, the model predicts the least singular agreement for *de-de* items and the most singular agreement for *het-het* items, with different-gender conjoined NPs (*de-het* and *het-de* items) in between. The effect of gender on number agreement that is predicted by the model is therefore one of morphophonological ambiguity. However, this contrasts with the results of our sentence completion task, in which the least singular agreement was found with the *het-de* items. Thus, there are important differences between the pattern of predicted results and the

results from our sentence completion task in terms of how gender affects number agreement.

Discussion

In Experiment 1, we investigated the role of grammatical gender on number agreement with conjoined NPs in Dutch. Through a corpus-based pre-study, we established that singular agreement with conjoined NPs occurs when the conjoined nouns are grammatically and notionally singular. In a sentence completion task, we elicited subject–verb agreement with conjoined NPs that had a range of notional valuations, and we manipulated the grammatical gender of the determiners on the singular nouns. In line with predictions, participants produced virtually no singular agreement with less notionally singular animate/count items. Further, within the mass/deverbal category, there was a significant effect of notionality, as items that were more notionally singular were more likely to take singular agreement than less notionally singular items.

Grammatical gender also impacted number agreement, over-and-above the notional effect. Conjoined NPs in which both nouns had the same gender exhibited more singular agreement than conjoined NPs with different gender. Furthermore, there was a marginally significant effect of the closest NP, because participants produced more singular verbs when the closest NP had an unambiguously singular determiner (*het*) than when the determiner on the closest NP was ambiguous for number (*de*).

We implemented a computational version of the Marking and Morphing model to compare the model's predictions to our experimental results. While the model reasonably approximated our notional effects, there were differences between the model's predictions and our experimental results in terms of grammatical gender. Specifically, the model predicted the least

singular agreement with *de-de* NPs. However, in the sentence completion task, we found the least singular agreement with *het-de* NPs. This lower rate of singular agreement with NPs that had mismatching gender is incompatible with a morphophonological account that is driven solely by ambiguity. This suggests that Marking and Morphing, with its reliance on ambiguity, is unable to account for the full range of morphophonological effects on agreement. Similarly, a Competition model cannot account for the morphophonological effects simply through distributional patterns, which would also rely on morphophonological ambiguity.

However, the effect of gender match is compatible with a cue-based retrieval model in which the gender cues on the determiner facilitate retrieval of both nouns in the conjoined NP. Within this account, the verb form is initially computed via another mechanism (such as Marking and Morphing or Competition), but if speakers feel uncertain at the point of producing the verb, they can go back and check to make sure they have established the correct agreement dependency with the subject head noun(s). If the determiners in the conjoined NP are distinct, this will aid the speaker in retrieving the full conjoined NP, and lead to more plural agreement.

In Experiment 2, we conduct a similar sentence completion study in German, a language that is closely related to Dutch, to test whether this effect of gender match will extend to another language. Furthermore, German has three grammatical genders and three gender-marked determiners, one of which is ambiguous for number. Therefore, as German has two gender-marked determiners that are unambiguous for number (*der* and *das*), this will enable us to manipulate same versus different grammatical gender with determiners that are unambiguous for number (*der*, which is used with singular masculine nouns and *das*, which is the determiner for singular neuter nouns), as well as same versus different gender with NPs involving an ambiguous determiner (*die,* which is used with feminine nouns and all plural nouns). In doing so, we can further test the conclusion from Experiment 1 that number ambiguity alone cannot account for the morphophonological effects of gender on number agreement.

Experiment 2: German

In Experiment 2, we manipulated grammatical gender with conjoined NPs in German, using items with a range of notional valuations, to investigate the impact of grammatical gender on number agreement. Because German has three determiners with grammatical gender (*der* and *das*, which are unambiguously singular,

and *die,* which is ambiguous for number), we can see whether gender mismatch leads to lower rates of singular agreement, even when both of the determiners are unambiguous for number. The predictions are similar to those for Experiment 1. Both Marking and Morphing and Competition models predict an effect of notionality, with more singular agreement when the conjoined NPs are notionally singular. These models also predict that, if there are morphophonological effects on number agreement, they are driven by number ambiguity. Therefore, both the Competition Model and Marking and Morphing would predict less singular agreement for items containing at least one determiner that is ambiguous for number (*die*), but that there would be no overall effect of gender match. However, if we find an effect of gender match versus mismatch, even when both determiners are unambiguously singular, this will provide additional evidence that ambiguity alone cannot account for morphophonological effects on agreement, and supports a cue-based retrieval model whereby mismatching gender on the two determiners would give the speaker additional cues that there are two distinct nouns within the subject NP.

Method

Participants
Thirty German native speakers in Germany completed the experiment (15 male; 15 female). The mean age of participants was 24.9 years (SD = 3.2; range: 18–30). Four participants reported acquiring another language prior to age six, but all considered German their most proficient language, so they were included in the analyses reported here.[2]

Materials
All experimental items consisted of a conjoined NP that contained two singular nouns along with their definite article, joined with the conjunction *und* "and". Items varied according to gender match (same vs. different gender) and notionality (animate/count vs. mass/deverbal). See Table 6 for sample items.

The final set of 120 experimental NPs was chosen from a larger pool of 206 candidate NPs based on the results of three norming tasks parallel to those from Experiment 1: a notionality rating task, a sensibility rating task and an article-use rating task. The 206 candidate items were divided into four lists of 51–52 items each. Seventy-three German native speakers, none of whom participated in the sentence completion task, rated a different list of experimental items, along with 11 filler items, for each task.

Table 6. Sample stimuli for the German sentence completion task (Experiment 2).

Gender condition		Animate/count nouns (Less notionally singular)	Mass/deverbal nouns (More notionally singular)
Different gender	Die-unamb.	Die Tasse und der Teller "The cup and the plate"	Die Reise und der Verkehr "The travel and the traffic"
	Unamb.-die	Das Pferd und die Kuh "The horse and the cow"	Das Leinen und die Seide "The linen and the silk"
	Unamb.-unamb.	Der Hut und das Hemd "The hat and the shirt"	Der Sand und das Gras "The sand and the grass"
Same gender	Die-die	Die Rübe und die Karotte "The beet and the carrot"	Die Ruhe und die Musik "The silence and the music"
	Der-der	Der Tisch und der Stuhl "The table and the chair"	Der Rauch und der Nebel "The smoke and the fog"
	Das-das	Das Huhn und das Küken "The chicken and the chick"	Das Heu und das Stroh "The hay and the straw"

For the critical experiment, 120 conjoined NPs were chosen so that there were 60 animate or count items, categorised as less notionally singular, and 60 mass or deverbal items, categorised as more notionally singular. Within each notional category, 30 items had the same grammatical gender, with an equal number of items containing two feminine, two masculine, or two neuter nouns ($n = 10$ each). Each notional category also contained 30 items with different grammatical gender. Among the animate/count items, 9 of these different grammatical gender items contained a feminine second noun, 11 items contained a feminine first noun and 10 items contained only masculine and neuter nouns (due to a coding error, one item was accidentally labelled as containing a feminine second noun when it in fact contained a feminine first noun, leading to the slight imbalance of feminine first versus second nouns in this condition). Among the mass/deverbal items, 10 different-gender items contained a feminine second noun, 10 items contained a feminine first noun and 10 items contained only masculine and neuter nouns. Across all experimental items, there was an approximately equal distribution of feminine ($n = 80$), masculine ($n = 78$) and neuter nouns ($n = 82$). See Table 7 for the descriptive results from the three norming tasks. ANOVAs treating notionality (animate/count vs. mass/deverbal) and grammatical gender match (same vs. different gender) as between-items variables revealed a main effect of notionality in the article-use ratings ($F(1, 116) = 71.12$, $p < .001$) because animate and count items were rated as more natural with both articles than mass and deverbal items, but no other effects or interactions in the article-use or sensibility ratings were significant (all $ps > .09$). There was a main effect of notionality in the notional rating task because animate and count items were rated as less notionally singular than mass and deverbal items ($F(1, 116) = 159.81$, $p < .001$), but there were no other significant effects or interactions (all $ps > .1$).

The 120 conjoined NPs were split into two experimental lists, such that participants heard 15 items in each condition (animate/count-different gender; animate/count-same gender; mass/deverbal-different gender; mass/deverbal-same gender), for a total of 60 experimental items. These experimental NPs were presented in a randomised order along with 100 filler NPs. As in Experiment 1, filler NPs included conjoined NPs that contained at least one plural noun, complex NPs that contained a variety of singular and plural head and local nouns, and simple singular and simple plural NPs. Across the experiment an equal number of NPs were – prescriptively speaking – grammatically plural and grammatically singular. All experimental and filler NPs were recorded by a female German native speaker.

Procedure
The procedure was identical to Experiment 1, except that participants were told to describe where things are or when they occur. They were also free to use any verb they wished. Stimulus presentation was controlled by Neurobehavioral Systems Presentation software (Version 16.5).

Scoring
Responses were transcribed and scored in the same manner as Experiment 1. Five percent of the data were scored by a second coder, with an interrater reliability of 99.5%. There were 1257 (69.8%) plural responses (696 with animate/count items; 561 with mass/deverbal items), 58 (3.2%) singular responses (6 with animate/count items; 52 with mass/deverbal items), and 485 (26.9%) miscellaneous errors (198 with animate/count items; 287 with mass/deverbal items). All responses, including miscellaneous items, are presented in Table 8.

Data analysis
As in Experiment 1, there was little variation in agreement patterns with the animate/count items, with few instances of singular agreement regardless of the grammatical gender of each NP (see Table 8 for descriptive results). Therefore, the primary statistical comparison

Table 7. Mean ratings (standard deviation) and range for notionality, sensibility and article use for German sentence completion task (Experiment 2).

	Notionality		Sensibility		Article use	
	M (SD)	range	M (SD)	range	M (SD)	range
Animate/count nouns						
Same gender	4.0 (0.44)	3.3–4.8	4.3 (0.36)	3.6–4.9	4.6 (0.45)	3.9–5.8
Different gender	3.9 (0.42)	3.3–4.7	4.2 (0.41)	3.4–4.9	4.7 (0.38)	4.0–5.7
Mass/deverbal nouns						
Same gender	2.8 (0.65)	1.6–3.8	4.2 (0.40)	3.2–4.8	3.9 (0.47)	3.0–4.9
Different gender	2.9 (0.33)	2.2–3.4	4.3 (0.38)	3.6–4.8	4.0 (0.49)	3.1–4.9

reported here included only the mass/deverbal items (see online supplemental material for the results of a preliminary analysis directly comparing animate/count items to mass/deverbal items). Notionality, treated as a continuous variable centred at the sample mean of 2.85, was entered as a fixed effect. Gender match (different vs. same gender) and closest NP (ambiguous-*die* vs. unambiguous-*der/das*) were entered as fixed effects, contrast-coded with the values −.5 and .5. We also entered the sensibility and article-use ratings as fixed effects, centred at the sample means of 4.20 and 3.95, respectively.

Results

In an initial model with the mass/deverbal items, we entered the interaction terms notionality × gender match, notionality × closest NP and gender match × closest NP. However, none of these interactions improved the overall fit of the model (all *p*s > .3), so they were not included in the final model reported here (see online supplemental material for the model with interaction effects). The random effect structure included random intercepts for items and participants,

and by-participant random slopes for closest NP. The inclusion of additional random slopes did not improve the model fit (all *p*s > .7); however, all results reported here also hold if the model includes the maximal uncorrelated random effect structure supported by the design.

As seen in Table 9, there were no significant effects of sensibility or article-use ratings. There was a significant effect of notionality because the proportion of singular responses increased on items rated as more notionally singular (i.e. items with lower notionality rating scores). There was a significant effect of gender match because the proportion of singular responses was higher for items with the same gender than items with different gender. There was no significant effect of ambiguity for the closest NP.

Discussion

In Experiment 2, we elicited subject–verb agreement with German conjoined NPs, capitalising on the flexibility of agreement with notionally singular items. We used the three gender-marked determiners in German (unambiguous *der/das* and ambiguous *die*) and manipulated

Table 8. Distribution of response type by scoring category and percentage of plural versus singular verbs for non-miscellaneous responses in German sentence completion task (Experiment 2).

Response type			
	Plural (% plural)	Singular (% singular)	Miscellaneous (misc. plural, misc. singular)
Animate/count nouns			
Different gender			
Die-unamb.	106 (100%)	0 (0%)	28 (23, 0)
Unamb.-die	121 (98%)	2 (2%)	43 (27, 2)
Unamb.-unamb.	120 (99%)	1 (1%)	29 (25, 1)
Same gender			
Die-die	115 (99%)	1 (1%)	34 (24, 0)
Der-der	113 (100%)	0 (0%)	37 (32, 0)
Das-das	121 (98%)	2 (2%)	27 (21, 1)
Mass/deverbal nouns			
Different gender			
Die-unamb.	102 (97%)	3 (3%)	45 (30, 1)
Unamb.-die	106 (96%)	4 (4%)	40 (25, 3)
Unamb.-unamb.	100 (97%)	3 (3%)	47 (33, 3)
Same gender			
Die-die	79 (91%)	8 (9%)	63 (36, 3)
Der-der	99 (94%)	6 (6%)	45 (35, 3)
Das-das	75 (73%)	28 (27%)	47 (18, 8)

Table 9. Summary of the mixed logit model for German mass/deverbal conjoined NPs (Experiment 2).

Predictor Fixed effects	Parameter estimates		Wald's test of individual coefficients		Δ(-2Λ)-test	
	Estimate	Std. error	z-Value	Pr (>|z|)	χ²	p
(Intercept)	−3.41	0.34	−9.99	<.001		
Sensibility	−0.34	0.66	−0.52	.602	0.28	.598
Article use	0.59	0.55	1.07	.29	1.17	.280
Notionality	−2.06	0.52	−3.99	<.001	15.26	.001
Gender match	1.70	0.54	3.13	.002	11.19	.011
Closest NP	−0.57	0.65	−0.87	.383	0.79	.851

same versus different grammatical gender within the conjoined NP, as well as the presence of morphophonological ambiguity. Both notionality and gender match affected agreement in German, but there was no effect of the ambiguity of the closest NP. Importantly, the effect of gender match was evident, even when both of the nouns were unambiguously singular; numerically, there was more singular agreement for all of the same-gender conditions in German than for any of the different-gender conditions. Numerically, there was also more singular agreement for the number-ambiguous *die-die* items than for the unambiguously singular *der-der* items. This provides additional evidence that the observed morphophonological effects are not solely driven by number ambiguity, but that gender match itself plays an important role, and suggests that models of agreement that only incorporate morphophonological effects through ambiguity are insufficient to account for all of the ways in which gender can affect number agreement.

Items with matching neuter determiners, *das-das*, had an especially high rate of singular agreement (27%), even compared to *der-der* (6%), which is also unambiguously singular. This parallels results from Experiment 1 with Dutch, where the *het-het* (neuter) determiners also showed increased singular agreement. This is likely because the nouns in this category in both languages were often deverbals (e.g. *Das Lachen und das Weinen* "the laughing and the crying"), which are often notionally singular. Indeed, amongst the same-gender items, the *das-das* items were rated as the most notionally singular on the notionality rating task (*M* = 2.02; SD = .31; range: 1.63–2.42). The other same-gender items were less notionally singular (*der-der: M* = 3.04; SD = .30; range: 2.68–3.59) (*die-die: M* = 3.35; SD = .32; range: 2.84–3.79). These differences in notionality likely explain the especially high rate of singular agreement among the *das-das* items in Experiment 2. The fact that these neuter determiners showed especially high rates of singular agreement in both German and Dutch highlights the value of cross-linguistic comparisons and

shows how important it is to consider that there may be inherent differences in notionality between different grammatical genders, due at least in part to the derivational principles that drive word formation in languages. Nevertheless, the fact that there was a significant effect of gender match, even when including notionality as an additional variable, highlights that such differences in notionality across genders cannot fully explain the present findings.

The Marking and Morphing model and the Competition model would both predict effects of notionality and morphophonological ambiguity. However, just as with Experiment 1, neither model would predict an effect of gender match. The predictions of Marking and Morphing, in terms of morphophonology, are as follows. The Marking and Morphing model predicts that every piece of unambiguously singular morphology, such as what is specified on determiners, increases the likelihood of singular agreement. Therefore, the model – and its corresponding implementation (see Experiment 1) – would predict that, given equal notionality ratings, conjoined NPs with two unambiguously singular determiners should have the most singular agreement, regardless of gender match and regardless of whether the language has a two-gender system like Dutch versus a three-gender system like German. The Competition model would make a similar prediction, as unambiguously singular determiners most often co-occur with singular verbs.

The effect of gender match is, however, predicted by a cue-based retrieval model, as the cues from each of the gender-marked determiners would facilitate retrieval of both NPs when the determiners have different-gender cues. The fact that German has two unambiguously singular determiners provides stronger evidence that this effect of gender match is not driven by morphophonological ambiguity, but that the mismatching gender features themselves are reducing the likelihood of singular verb agreement. This hypothesis will be discussed in greater detail in the General Discussion.

General discussion

Overall, this set of experiments provides important insight into the role of morphophonology on agreement with conjoined NPs in both Dutch and German. In Experiments 1 and 2, there was a significant effect of notionality, as participants produced more singular agreement when the conjoined NPs were notionally singular. There was also a significant effect of gender match, as speakers produced more singular agreement when both nouns in the conjoined NP had the same gender. Furthermore, Experiment 1 (Dutch) showed a marginally

significant increase in the rate of singular verb agreement when the closest NP had an unambiguously singular determiner, while Experiment 2 (German) showed no effect of the ambiguity of the determiner on the closest NP.

The notional effect is consistent with both the Marking and Morphing model (Eberhard et al., 2005) and the Competition model (Mirković & MacDonald, 2013). However, neither the Marking and Morphing model nor the Competition model predicts the effect of gender match. To complement the experimental data, we performed a computational implementation of Marking and Morphing in Dutch, which showed an important difference between the results of Experiment 1 and the model's predictions; specifically, the Marking and Morphing model predicts the most plural agreement when both determiners are ambiguous for number (*de-de*), but it does not predict the effect of gender match that we saw in our experimental data.

The role of morphophonological ambiguity in agreement

To date, morphophonological ambiguity has been seen as the primary driver of morphophonological effects in agreement, in both the Marking and Morphing (Eberhard et al., 2005) and Competition (Mirković & MacDonald, 2013) models, and previous work testing the tenets of these models (e.g. Antón-Méndez & Hartsuiker, 2010; Badecker & Kuminiak, 2007; Franck et al., 2008; Hartsuiker et al., 2003; Mirković & MacDonald, 2013) has shown that morphophonological ambiguity can play an important role in agreement. In Experiment 1, we did observe numerically less singular agreement with the ambiguous determiners *de-de* than with the unambiguously singular *het-het* determiners, and there was a marginally significant effect of the closest NP, such that when the second NP had an unambiguously singular *het*-noun, speakers produced more singular agreement than when the second noun had the ambiguous determiner *de*. However, in Experiment 2 (German), the effect of the ambiguity of the closest NP was not significant, and the unambiguously singular *der-der* items showed numerically less singular agreement than the number-ambiguous *die-die* items. Altogether, this suggests that, while morphophonological ambiguity may have a role in agreement, it is not the primary driver of gender effects on number agreement with conjoined NPs.

Same versus different gender

More critical for determining the extent to which current models of agreement can account for the present findings is the result that same-gender conjoined NPs exhibited higher rates of singular agreement than different-gender conjoined NPs in both languages and that this effect was significant, even in a statistical model that took notionality into account. This contrasts with the predictions of the Marking and Morphing model (Eberhard et al., 2005). The Competition model (Mirković & MacDonald, 2013), in which agreement is driven purely by co-occurrence patterns, such that singular verbs are activated whenever unambiguously singular determiners are used, and plural verbs whenever ambiguous determiners are used, is also unable to account for the present effects of gender match on number agreement.

However, it is possible that the effects of gender match could be explained with cue-based retrieval, similar to the effect of gender match on agreement attraction in Spanish (Lorimor et al., 2015), as speakers could be keeping track of the gender of the nouns, which they would use to help retrieve the subject at the point they are planning the verb. In such a scenario, if the two nouns have different genders, this would make it easier for speakers to keep track of the fact that there were two distinct nouns within the subject NP, leading to less singular agreement in the gender mismatch condition, compared with the gender match condition.

Conclusion

Overall, the results of our experiments provide evidence that agreement with conjoined NPs in Dutch and German is influenced by both notional number and morphophonology and that the effects of morphophonology are not primarily driven by morphophonological ambiguity, which is the main vehicle for morphophonological effects in both Marking and Morphing (Eberhard et al., 2005) and Competition models (Mirković & MacDonald, 2013). However, there is mounting evidence that cue-based retrieval plays an important role, both in agreement comprehension (Lago et al., 2015; Tanner et al., 2014; Wagers et al., 2009), and in agreement production (Lorimor et al., 2015). Cue-based retrieval can account for the effect of gender match that was observed in both Experiment 1 and Experiment 2, since the distinct gender features on each determiner can facilitate retrieval of the full subject NP. Finally, the present experiments highlight how, moving forward, production research must continue to include a wider variety of languages and linguistic constructions, and to consider how research on language comprehension can inform language production, as only then can we begin to understand the precise mechanisms that underlie agreement processes during language production.

Notes

1. Eberhard et al. (2005) define $S(m)$ as the number Specification * the contrastive frequency (C_{freq}), which is calculated by taking the log(10) of the total frequency over the log of the plural frequency. We do not use C_{freq} here, because for all nouns that have a plural counterpart, their Specification is 0, thus cancelling out the C_{freq} term. For the other items – the invariant singulars – C_{freq} terms would be undefined, as the log of zero is undefined.
2. A second analysis excluding these four participants revealed the same pattern of significant effects as the results from all 30 participants.

Acknowledgements

Disclosure statement

Funding

References

Adani, F., van der Lely, H. K. J., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua, 120*, 2148–2166. doi:10.1016/j.lingua.2010.03.018

Antón-Méndez, I., & Hartsuiker, R. (2010). Morphophonological and conceptual effects on Dutch subject-verb agreement. *Language and Cognitive Processes, 25*, 728–748. doi:10.1080/01690960903513818

Antón-Méndez, I., Nicol, J., & Garrett, M. F. (2002). The relation between gender and number agreement processing. *Syntax, 5*, 1–25. doi:10.1111/1467-9612.00045

Badecker, W. (2007). A feature principle for partial agreement. *Lingua, 117*, 1541–1565. doi:10.1016/j.lingua.2006.06.006

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language, 56*, 65–85. doi:10.1016/j.jml.2006.08.004

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua, 122*, 1053–1069. doi:10.1016/j.lingua.2012.02.007

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45–93. doi:10.1016/0010-0285(91)90003-7

Corbett, G. G. (2006). *Agreement*. Cambridge: Cambridge University Press.

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36*, 147–164. doi:10.1006/jmla.1996.2484

Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language, 41*, 560–578. doi:10.1006/jmla.1999.2662

Eberhard, K. M., Cutting, J. C., & Bock, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531–559. doi:10.1037/0033-295X.112.3.531

Foote, R., & Bock, J. K. (2012). The role of morphology in subject-verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes, 27*, 429–461. doi:10.1080/01690965.2010.550166

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes, 23*, 329–374. doi:10.1080/01690960701467993

Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory and Cognition, 31*, 1316–1326. doi:10.3758/BF03195814

Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language, 48*, 760–778. doi:10.1016/S0749-596X(03)00010-X

Haskell, T. R., Thornton, R., & MacDonald, M. C. (2010). Experience and grammatical agreement: Statistical learning shapes number agreement production. *Cognition, 114*, 151–164. doi:10.1016/j.cognition.2009.08.017

Hoekstra, H., Moortget, M., Renmans, B., Schouppe, M., Schuurman, I., & van der Wouden, T. (2003). *CGN syntactische annotatie*. Retrieved from http://lands.let.ru.nl/cgn/doc_Dutch/topics/version_1.0/annot/syntax/info.htm

Humphreys, K. R., & Bock, J. K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review, 12*, 689–695. doi:10.3758/BF03196759

Jaeger, F. T. (2008). Categorical data analyses: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. doi:10.1016/j.jml.2007.11.007

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language, 82*, 133–149. doi:10.1016/j.jml.2015.02.002

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 375–419. doi:10.1207/s15516709cog0000_25

Lorimor, H. (2007). *Conjunctions and grammatical agreement* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign, Champaign, IL..

Lorimor, H., Bock, J. K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes, 23*, 769–799. doi:10.1080/01690960701774182

Lorimor, H., Jackson, C. N., & Foote, R. (2015). How gender affects number: Cue-based retrieval in agreement production. *Language, Cognition and Neuroscience, 30*, 947–954. doi:10.1080/23273798.2015.1047461

Mirković, J., & MacDonald, M. C. (2013). When singular and plural are both grammatical: Semantic and morphophonological effects in agreement. *Journal of Memory and Language*, *69*, 277–298. doi:10.1016/j.jml.2013.05.001

Morgan, J. (1984). Some problems of determination in English number agreement. In G. Alvarez, B. Brodie, & T. McCoy (Eds.), *Proceedings of the first eastern states conference on linguistics* (pp. 69–78). Columbus: Ohio State University, Department of Linguistics.

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago, IL: University of Chicago Press.

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*, 413–425. doi:10.1016/j.jml.2008.02.002.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools.

Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanism and asymmetrical attraction. *Journal of Memory and Language*, *76*, 195–215. doi:10.1016/j.jml.2014.07.003

Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, *48*, 740–759. doi:10.1016/S0749-596X(03)00003-2

van Koppen, M. (2005). *One probe - two goals: Aspects of agreement in Dutch dialects* (PhD dissertation, LOT-Publications No. 105). Leiden University, Leiden.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, *128*, 442–472. doi:10.1037/0033-2909.128.3.442

Vigliocco, G, Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, *11*, 407–442. doi:10.1080/016909696387169

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206–237. doi:10.1016/j.jml.2009.04.002