"Norwegian agreement clashes on the football field"
Laura A. Janda and Tore Nesset
UiT The Arctic University of Norway

Kibrik (2019) argues that a cognition-to-form approach to agreement is superior to the traditional form-to-form approach that is fraught with problems. We concur that it makes more sense to examine agreement from the perspective of cognitive representation and present a small study of how the semantics of adjectives contributes to the use of singular vs. plural agreement with Norwegian collective nouns naming football teams.

In the Norwegian Web As Corpus (NoWaC: https://tekstlab.uio.no/glossa2/nowac_1_1), we find these examples, both of which report that the football team named Glimt has qualified to play in the Premier League:

1a) (NoWaCID:33997607)

| *Glimt* | *er* | ***klar*** | *for* | *Tippeliga-en* |
| [Glimt | be.PRS[1] | **ready.SG** | for | premier.league-DEF] |

'Glimt is **ready.SG** for the Premier League'

1b) (NoWaCID:4090386)

| *Glimt* | *er* | ***klar-e*** | *for* | *Tippeliga-en* |
| [Glimt | be.PRS | **ready-PL** | for | premier.league-DEF] |

'Glimt are **ready.PL** for the Premier League'

The examples are identical except that 1b) has a plural marker on the predicate adjective, wheras 1a) indicates singular. The name of a football team is a collective term that can refer to a single entity (an organization) or to the players. We argue that conceptualization as a single entity motivates singular agreement, whereas conceptual reference to the players motivates plural agreement.

Variation in number agreement with collective nouns is well-known both cross-linguistically (cf. Corbett 1983 and 2006 about Russian, Hagåsen 2019 about Swedish, Levin 2001 about English), and with respect to Norwegian (cf. Faarlund et al. 1997, Åfarli and Vangsnes 2021). Nesset et al. 2022 is the first study to investigate this singular-plural variation in Norwegian through a combination of factors on the basis of corpus data, finding that the following are significant: the category of collective noun (political party, country, police, etc.), the distance between the subject and the predicate adjective, and whether the agreement target is an adjective or a participle.

We continue this line of research with a new dataset and with special focus on the role of adjectival semantics in determining preference for singular vs. plural agreement. We have collected sentences from NoWaC with the names of Norwegian football team names as subjects: *(Bodø) Glimt*, *Brann*, *HamKam*, *Hødd*, *Jerv*, *Rosenborg*, *Strømsgodset*, and *Viking*. Exclusive focus on football teams eliminates variation caused by differences across categories of collectives, and a chi-squared test shows that there is no significant variation

---

[1] Note that person and number are not distinguished on the verb in Norwegian.

in number agreement across the football teams. Subject-predicative distance was held to a minimum by collecting only sentences in which the predicative adjective followed immediately after the verb or was modified by an adverb; no examples with prepositional phrases or other intervening words were included and the distances that were observed were found to be statistically insignificant for choice of singular vs. plural. The data was manually cleaned, removing sentences with adjectives that cannot show the singular vs. plural contrast (for example, comparatives and superlatives), and sentences in which the subject was not a football team: for example, *jerv* means 'wolverine', an animal found in abundance in Norwegian forests as well as sentences. The remaining 495 examples were manually tagged for the number of the predicate adjective (singular vs. plural), the type of adjective (adjective vs. participle), and the lemma of the adjective.[2]

Our goal was to investigate the role of the adjective in the choice between singular vs. plural agreement by means of a statistical model. Of the 495 examples in our database, 291 had adjectives marked for plural, and the remaining 204 adjectives were singular. This means that the baseline for our study was 59%, namely the accuracy achieved if one always guessed plural. A statistical model must be evaluated in comparison with this baseline.

We proposed the following generalized linear mixed-effects model to capture the relationship between Number (singular vs. plural marking on the adjective) as the result variable, and the type of adjective and adjective lemma as predictor variables, following this formula in R:

Formula: Number ~ 1 + Adjective_vs_participle + (1 | Adjective_lemma)

Fixed effects:				
	Estimate Std.	Error	z value	Pr(>|z|)
(Intercept)	0.9767	0.2484	3.932	8.41e-05 ***
Adjective_vs_participle: participle	-3.5822	0.8013	-4.470	7.81e-06 ***

Table 1: Fixed effects results for mixed effects logistic regression model

The Intercept is the prediction of plural (positive value) when the adjective is not a participle, and the prediction is of singular (negative value) when the adjective is a participle. This model performs quite well, with a C score of 0.89 (a value above 0.8 is considered good or excellent, cf. Gries 2021: 335–336), and conditional R-squared values ranging from 0.46 to 0.52. The model correctly predicts the number of the adjective for 406 of the 495 examples, giving a classification accuracy of 82%. This accuracy is well above the baseline of 59%, as are all measures of Precision and Recall.[3] A statistical test shows that the probability that a model could do this well by chance is 1.48035e-28 (virtually zero).

The model predictions for the status of the adjective are visualized in Figure 1. The y-axis indicates the predicted probability of plural, and a dotted line shows the 50% line above which plural is predicted. The brackets on the bars indicate the 95% confidence interval. The

---

[2] Examples were also tagged for the name of the football team, the subject-predicative distance, and tense (past vs. present), but none of these factors were found to be significant.
[3] Precision for plural = 83%, Recall for plural = 87%, Precision for singular = 80%, Recall for singular = 75%.

width of the bars indicates the relative proportion of data for the two types. There were only 48 examples with participles, and participles clearly prefer singular agreement, although there are 5 examples with plural.
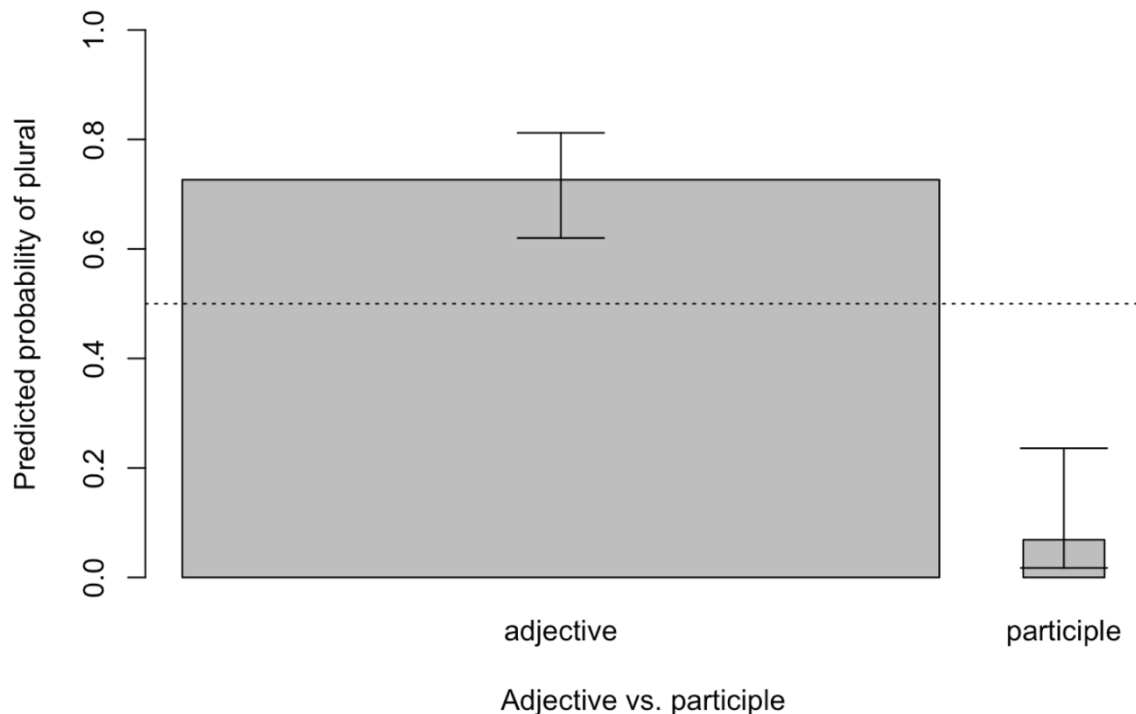

Figure 1: model predictions for the status of the adjective

Crucially, however, the fixed effect of the status of the adjective cannot on its own provide us with a good statistical model. On its own the status of the adjective yields only 66% classification accuracy and has poor scores (C = 0.597; R-squared = 0.141), and of course it tells us little about the 447 examples that do not contain participles. We achieve a viable model for this dataset only when we take into account the individual preferences of the adjectives as a random effect. This means that it is the adjectives themselves, and therefore their semantics, that play the major role.

We argue that the semantics and conventional uses of adjectives motivate number agreement. More precisely, among adjectives we observe a continuum that ranges from a preference for conceptualization of the referent as a number of individuals and thus plural, to a preference for conceptualization of the referent as whole and thus singular. Fortunately for this study, sports reporting is a relatively colloquial and hackneyed style, which means that we get insight into spoken norms across a compact selection of adjectives. Our 495 examples contain in sum only 138 unique lemmas, fourteen of which occur more than ten times: thirteen of these are adjectives, and one is a participle: *ubeseiret* 'undefeated'. The model predictions for the random effects and the observed number distributions for the adjectives are gathered in Table 2. Table 2 is ordered according to the random effects values that measure preferences against the 59% plural baseline: positive values indicate a further plural preference, negative values indicate an increased preference for singular.

Adjective lemma	Random effects value	Pural/singular observations
effektiv 'effective'	1.70	11/0
tafatt 'helpless'	1.64	10/0
god 'good'	0.76	36/6
sjanseløs 'hopeless'	0.51	10/2
farlig 'dangerous'	0.42	13/3
heldig 'lucky'	0.02	19/7
villig 'willing'	-0.73	7/6
klar 'ready'	-0.91	11/11
dårlig 'bad'	-1.15	10/13
enig 'agreed'	-1.27	8/12
nær 'near'	-1.73	6/15
ferdig 'done for'	-1.87	3/10
avhengig 'dependent on'	-2.12	3/13

Table 2: Random effects and observed values for adjectives with >10 attestations

Examination of examples from the extreme ends of the distribution in Table 2 supports our claim that some adjectives motivate conceptualization of the teams as groups of individuals, while others focus more on the team as a single whole. In 2) it is clear that the speaker is talking about the Viking players, in contrast to the Rosenborg players, while 3) is about the status of the team organization that is dependent on its sponsors.

2) (NoWaCID:8886729)
*Viking er **tafatte.PL** og gir stadig ballen fra seg til langt mer aggressive Rosenborg-spillere.*
'Viking are **helpless.PL** and are always letting the ball go to the much more aggressive Rosenborg players.'

3) (NoWaCID:24423445)
*Brann er **avhengig.SG** av sponsorer for å ha sjansen til å levere sportslige resultater.*
'Brann is **dependent.SG** on sponsors to have a chance to deliver competitive results.'

In the middle of the distribution, we find adjectives that are compatible with both a plural, as in 4), and with a singular, as in 5) conceptualization of the referent.

4) (NoWaCID:32736864)
*Glimt var litt **heldige.PL** som fikk gå til pause på stillingen 2 – 2*
'Glimt were a little **lucky.PL** that they got to go to halftime when the score was 2 – 2'

5) (NoWaCID:7035482)
*Strømsgodset er **heldig.SG** som leder på Ullevaal.*
'Strømsgodset is **lucky.SG** to be the leader at Ullevaal.'

Oddly, *god* 'good' has a preference for plural as in 6), while the antonym *dårlig* 'bad' has a preference for singular, as in 7). We speculate that there may be some cultural factor at play

here: perhaps it is less acceptable in Norway to call out the players as individuals as being 'bad', while it is normal to praise them as 'good'.

6) (NoWaCID:4778489)
*Brann er **gode.PL** når de får til sitt spill.*
'Brann are **good.PL** when they succeed in their plays.'

7) (NoWaCID:19582531)
*Brann er **dårlig.SG** på dødball.*
'Brann is **bad.SG** in a deadball situation.'

We have shown that number agreement for Norwegian predicate adjectives noun is sensitive to the semantics of lemmas and that this is a scalar rather than categorical phenomenon. More precisely, adjectives vary in their compatibility with a conceptualization of a team as a group of players as opposed to as a single unit. These findings conform to Kibrik's (2019) idea that the cognition-to-form approach is a superior description of agreement.

References

Corbett G.G. 1983. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic.* London: Croom Helm.

Corbett G.G. 2006. *Agreement.* Cambridge: Cambridge University Press.

Faarlund J. T., Lie S., Vannebo K. I. 1997. *Norsk referansegrammatikk.* Oslo: Universitetsforlag.

Gries S. Th. 2021. *Statistics for linguistics with R*, 3rd rev. & ext. edn. Berlin: De Gruyter.

Hagåsen L. 1992. Sverige var intresserad och facket ska vara obunden – en typ av neutruminkongruens. In: G. Widmark et al. (eds.). *Svenska i harmoni*. Uppsala: Hallgren & Fallgren, 60–75.

Kibrik A. A. 2019. Rethinking agreement: Cognition-to-form mapping. *Cognitive linguistics*, v. 30, n. 1, 37–83.

Levin M. 2001. *Agreement with collective nouns in English*. Lund: Lund University Press.

Nesset T., Enger H. O., Janda L. A. 2022. Er politiet sikker eller sikre? Adjektivkongruens ved kollektiver i norsk. *Arkiv för nordisk filologi* v. 137, 123–144.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Åfarli T. A., Vangsnes Ø. A. 2021. Formell og semantisk adjektivkongruens i norsk. In: K. Hagen, A. Hjelde, K. Stjernholm, Ø.A. Vangsnes (eds.). *Bauta: Janne Bondi Johannessen in memoriam*, *Oslo Studies in Language* v. 11, n. 2, 527–540.