

Morphological and semantic agreement beyond hybrid nouns*

Zheng Shen & Peter W. Smith

Goethe University Frankfurt

1. Introduction

Cross-linguistically, certain nouns can optionally control two types of agreement, one agreement that reflects the semantic content of the noun, and the other that more closely reflects the morphological shape. These nouns have been termed HYBRID NOUNS, since they seem to be mixtures of separate strands of information that coexist on one and the same noun. For example, collective nouns like *committee* can optionally control singular and plural agreement in certain varieties of English in (1).

- (1) a. This committee has gathered.
b. This committee have gathered. (Corbett 1979, 2006, Smith 2015)

Collective nouns are hybrid in the sense that they are unambiguously morphologically singular in the above examples (as they all have plural variants), but they have the semantics of a group of entities (committees etc. must generally be formed of more than one member). Hybrid noun agreement is not limited to collective nouns or to English. The word for ‘doctor’ in Russian *vrač* is grammatically masculine, but can refer to female doctors. As is seen in (2), when a female doctor is referred to, both the adjective and the verb can show masculine and feminine agreement. (Corbett 1979).

- (2) a. Novyj vrač skazal.
 new.MASC doctor spoke.MASC
b. Novaja vrač skazala.
 new.FEM doctor spoke.FEM
 ‘A new (female) doctor spoke.’

*Thanks to participants and reviewers of NELS 49 and the workshop on Multiple Agreement across Domains 2018, ZAS Berlin, as well as audiences at Goethe University Frankfurt and Rutgers University. The research is supported by the DFG project *A General Theory of Multi-valuation* (SM: 514/1-1). This paper is an extension of section 4.2 and 4.3 of Shen (2019).

In Chichewa the word for ‘hero’, *ngwazi*, belongs to class 9. Corbett (1991) observes that it can also control class 1 agreement which is the default animate class as is shown in (3).¹

- (3) a. ngwazi y-athu y-oyamba
 hero 9-our 9-first
 b. ngwazi w-athu w-oyamba
 hero 1-our 1-first
 ‘our first hero’

The existing analyses for hybrid noun agreement assume two distinct sets of features on hybrid nouns available to control agreement (Wechsler and Zlatič 2000, Smith 2015, 2017). One set is the morphological (CONCORD/uninterpretable features, uF), the other is the semantic features (INDEX/interpretable features, iF).² The alternating agreement patterns result from the targets agreeing with the respective set of features on the hybrid noun. For example, the singular agreement pattern with English collective nouns in (4a) results from the verb agreeing with the singular uF:[SG]. The plural agreement pattern in (4b) results from the verb agreeing with the plural iF:[PL]. For the rest of the paper, we will refer to agreeing with iF as semantic agreement and agreeing with uF as morphological agreement, following previous research, particularly (Smith 2015, 2017).

- (4) a. The committee_{uF:[SG],iF:[PL]} has_[SG] gathered.
 b. The committee_{uF:[SG],iF:[PL]} have_[PL] gathered.

Assuming the dual feature analysis for the hybrid nouns is on the right track, one obvious question arises: what about the non-hybrid nouns like *student*? Two plausible hypotheses one may entertain are (i) unlike hybrid nouns, the non-hybrid nouns only possess one set of features; (ii) like hybrid nouns, the non-hybrid nouns also possess both iF and uF. These two hypotheses are equally possible and teasing them apart is an empirical issue, however finding the relevant empirical evidence can be difficult. The canonical agreement configuration would not be helpful because the predictions from both hypotheses give the same result. If non-hybrid nouns only have one set of feature, the verb would show the value of that feature as in (5). If non-hybrid nouns have two sets of features, they would have the same value ([SG] in (6)), the verb would show singular, no matter which feature actually controls the agreement.

- (5) The student_{F:[SG]} has_[SG] arrived.

- (6) a. The student_{uF:[SG],iF:[SG]} has_[SG] arrived.

¹See also Puškar-Gallien (2019) for data of the polite pronoun in Ukrainian.

²A slight distinction arises between existing theories in how directly responsible these features are for morphological form or semantic interpretation. For Smith (2015) the uFs and iFs are directly responsible, however, in Wechsler and Zlatič’s system, the CONCORD and INDEX features are simply closely related to the ultimate interface features. Since the distinction doesn’t matter any further here, we will not discuss it in any more detail, and we assume Smith’s analysis.

- b. The student_{uF:[SG],iF:[SG]} has_[SG] arrived.

In this paper, we will argue that in certain non-canonical agreement configuration, namely MULTI-VALUATION, the agreement patterns differ when a target agrees with the iF from when it agrees with uF, thus providing evidence supporting the dual feature analysis for non-hybrid nouns.

2. Multi-valuation in TP right-node-raising

Multi-valuation refers to cases where one agreement target agrees with multiple agreement controllers. Yatabe (2003), Kluck (2009), Grosz (2015), Shen (2017a,b, 2018, 2019), Belk and Neeleman (2018) discuss TP right node raising (TP RNR) in (7) as a case of multi-valuation.

- (7) John is proud that Maria_[SG], and Bill is glad that Eva_[SG], **have**_[PL]/**has**_[SG] traveled to China.

In (7), the embedded predicate *has/have traveled to China* is shared by two conjoined clauses. The auxiliary can either be singular or plural. There is no semantic difference based on the choice of the agreement. We will label the singular agreement as distributive agreement since the [SG] is evenly distributed on the subjects and the auxiliary. We will label the plural agreement in (7) as summative agreement since on the surface, the two [SG] values get summed up on the auxiliary into a [PL] value.

Both distributive and summative agreement are observed in English, as well as in Italian, French, Slovak, and some varieties of German. On the other hand, languages including Slovenian, Dutch, Finnish, Greek, Polish, and Serbo-Croatian only allow distributive agreement but not summative agreement, as is shown for Slovenian in (8). Meanwhile, in Russian, only the summative agreement is allowed as is shown in (9). See Grosz (2015), Shen (2018, 2019) for more cross-linguistic data on TP RNR.

- (8) Jure misli da Maja_[SG], in Boris verjame da Sara_[SG], **potuje**/***potujeta** na
Jure thinks that Maja, and Boris believes that Sara travel.SG/*DL to
kitajsko.
China
'Jure thinks that Maja, and Boris believes that Sara travels to China.' (Slovenian)

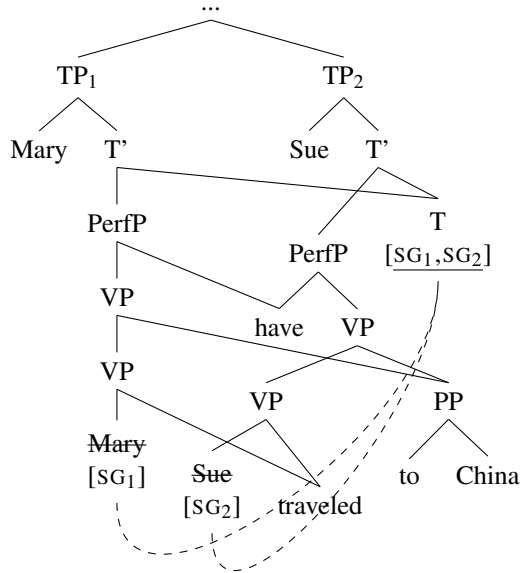
- (9) Ivan думаet chto Masha_[SG], a Vasya думаet chto Dasha_[SG],
Ivan thinks comp Masha, and Vasya thinks COMP Dasha,
ezdil-i^{??} **ezdil-a** v Kitaj.
went-PL^{??} went-SG to China
'Ivan thinks that Masha, and Vasya thinks that Dasha, went to China.' (Russian)

Having established the cross-linguistic facts on TP RNR, we discuss the assumption that both agreement patterns involve multi-valuation. It is worth mentioning that what's relevant

for the purpose of this paper is that the verb/auxiliary in TP RNR is in an agreement relationship with both subjects, i.e. multi-valued. The specific implementation of this property is an important research question but not immediately relevant.

Two potential analyses for TP RNR were discussed in Grosz (2015). One involves a multi-dominance analysis where the predicate is *structurally* shared by the two clauses as is shown in (10). According to this analysis, each node in the predicate is simultaneously dominated by elements in each clause respectively. See Kluck (2009), Grosz (2015), Shen (2019) for evidence supporting the multi-dominance analysis.

(10) John is proud that Mary, and Bill is glad that Sue, **has/have** traveled to China.



A direct consequence of the multi-dominance analysis is that the T head gets two singular feature values from the embedded subjects, *Mary* and *Sue*. Note that *Mary* and *Sue* themselves are not conjoined. In the multi-dominance approach for TP RNR, both distributive and summative agreement on T result from multi-valuation.

An alternative approach for TP RNR is a hybrid approach: summative agreement results from multi-dominance as is discussed above, while distributive agreement results from ellipsis, where the predicate in the first clause gets deleted at PF, given the presence of an identical predicate in the second clause as is shown in (11). Under the ellipsis analysis, the ‘shared’ predicate in distributive agreement cases belongs to the second clause only.

(11) John is proud that Mary ~~has traveled to China~~, and Bill is glad that Sue has traveled to China.

Kluck (2009), Larson (2012), Shen (2019) show convincingly that the ellipsis analysis shown in (10) cannot be maintained. The readers are referred to these papers for the detailed arguments. For the rest of the paper we will follow these arguments and proceed under the assumption that both distributive and summative agreement in TP RNR result from multi-dominance and thus involve multi-valuation.

3. Morphological and semantic agreement in multi-valuation

Having established the basic facts on multi-valuation, we argue that the alternating agreement patterns observed in multi-valuation result from morphological agreement and semantic agreement with non-hybrid nouns. The arguments come from a set of similarities between, on the one hand, distributive agreement and morphological agreement, and on the other, summative agreement and semantic agreement. This conclusion will in turn support the claim that even non-hybrid nouns have both uF and iF.

3.1 Distributive agreement results from morphological agreement

Morphological agreement in hybrid noun agreement is sensitive to the morphological marking on the agreement controller, rather than the semantic referent. In (12), the collective predicate *gathered* requires the subject to be semantically plural: **John has gathered*. Nevertheless, the singular auxiliary *has* makes reference to the singular morphological marking on the subject.

(12) This committee has gathered.

Distributive agreement in multi-valuation shares this property. In (13), the sentence refers to a scenario where both the student and the teacher have traveled to China (given the factive matrix predicates: *is proud/glad that...*). At the same time, the morphological markings on both of the subjects are singular. The singular marking on the multi-valued *has* in (13) is sensitive to the morphological marking and not the semantic reference. This is expected if distributive agreement results from morphological agreement.

(13) John is proud that the student_[SG], and Bill is glad that the teacher_[SG], **has**_[SG] travelled to China.

Another argument comes from mismatch cases. It is known that when a verb receives two values that do not match, it is often problematic for the verb to determine which agreement value to show (Zaenen and Karttunen 1984, Pullum and Zwicky 1986). Bhatt and Walkow (2013) observe that when one verb agrees with two mismatching objects in Hindi-Urdu, the verb shows closest conjunct agreement (CCA), i.e. agreeing with the linearly closest object. We will follow their analysis that CCA results from agreement being delayed in the PF interface after linearization, see Arregi and Nevins (2012), Smith (2015, 2017) for independent evidence. Moreover, Corbett (1979) proposes that CCA is a resolution of mismatches in morphological agreement. If distributive agreement is in essence morphological agreement, one would expect CCA to be a viable repair strategy. This prediction is borne out: Kluck (2009) uses experimental data to show that Dutch only allows distributive agreement in TP RNR and when the two embedded subjects have mismatching values, the shared verb agrees with the 2nd/linearly closest subject as is shown in (14) (Kluck 2009, 159).

- (14) Anna beweerde dat wij, maar Steven zei dat jij, het gas aan **had/*hadden**
Anna claimed that 1PL, but Steven said that 2SG, the gas on **have.SG/*.PL**
laten staan.
let stand
'Anna claimed that we, but Steven said that you left the gas on.' (Dutch)

3.2 Summative agreement results from semantic agreement

Semantic agreement in hybrid noun agreement is sensitive to the semantic referent and not the morphological marking on the agreement controller. In (15), the plural marking is used despite the singular marking on the subject, since the reference of *the committee* includes the multiple members of the committee.

- (15) The committee have gathered.

Summative agreement shows the same property. In (16), neither embedded subject is plural, however, the auxiliary can be plural, since multiple individuals traveled to China.

- (16) John is proud that the student_[SG], and Bill is glad that the teacher_[SG], **have**_[PL] traveled to China.

Yatabe (2003) observes that when the embedded subjects have the same singular referent, plural agreement is ruled out in (17). This fact is not surprising if summative agreement makes reference to the semantic referent. Given that the two subjects refer to the same individual, the semantic agreement would also result in singular agreement.

- (17) The pilot claimed that the nurse from the United States₁, and, the sailor also claimed that the nurse from the United States₁, **was a spy/*were spies**.
(Yatabe 2003, 8)

Moreover, Yatabe (2003) claims that under disjunction, the plural agreement is ruled out (18). Belk and Neeleman (2018) point out that even under conjunction, the plural agreement is ruled out when it is clear that only one individual is referred to in the context in (19). It is obvious from these data that the summative agreement is connected to semantic agreement.

- (18) The pilot claimed that the first nurse, or the sailor proved that the second nurse **was a spy/?were spies**. (Yatabe 2003, 7)
- (19) [Someone has stolen the backdoor key, but we do not know who.]
John said that Bill, and Bill said that John, **has/*have** stolen the backdoor key.
(Belk and Neeleman 2018, 49a)

The final piece of evidence comes from mismatch cases. Recall that in Dutch where only morphological/distributive agreement is allowed, the mismatch case is resolved by CCA. In English where summative agreement is allowed, mismatching subjects will control plural agreement regardless of the order they come in. As is shown in (20), both orders of the embedded subjects result in plural agreement. This pattern is expected if summative agreement is agreeing with *iF*: the plural marking appears whenever multiple individuals have been to China. CCA, which is the repair strategy available for morphological agreement, is not available here.³

- (20) a. John is proud that $Mary_{iF:[SG]}$, and Bill is glad that the $twins_{iF:[PL]}$, **have**_[PL] been to China.
b. Bill is glad that the $twins_{iF:[PL]}$, and John is proud that $Mary_{iF:[SG]}$, **have**_[PL] been to China.

3.3 Summary: a dual feature analysis for non-hybrid nouns

The arguments laid out above indicate that distributive agreement results from morphological agreement and that summative agreement results from semantic agreement. Given that the relevant agreement controllers in TP RNR are non-hybrid nouns, the alternating agreement patterns (distributive and summative) support the dual feature analysis for all nouns: both hybrid and non-hybrid nouns have *iF* and *uF*. This is to our knowledge the first empirical evidence on this issue.

A note on the scope of the claim is in order. The arguments presented above apply to languages where both distributive and summative agreement are allowed, including English, Italian, French, Slovak, etc. Regarding languages that only allow distributive agreement, including Dutch, Finnish, Greek, Polish, Serbo-Croatian, and Slovenian, it is in principle possible that the non-hybrid nouns only have one set of features. It is also possible that non-hybrid nouns in these languages have both *iF* and *uF* as well, and for some reason only *uF* agreement is available in TP RNR. It is clearly the null hypothesis that languages behave alike, but further study is required to tease these two possibilities apart.

4. Consequences: nature of *uF* and *iF*

Given the discussion so far, multi-valuation can be modeled as in (21): a target agreeing with multiple *iF*:[SG] will be spelled out as plural and a target agreeing with multiple *uF*:[SG] will be spelled out as singular.

³It is curious that even though English allows morphological agreement in TP RNR, CCA is not available in mismatch cases as is shown in (i). Although the second embedded subject *Mary* is singular, the shared auxiliary cannot be singular according to our consultants. One hypothesis would be that CCA is a repair strategy which can only be activated as a last resort. Since English additionally allows semantic agreement in TP RNR, the repair strategy is blocked.

- (i) ??Bill is s glad that the $twins_{iF:[PL]}$, and John is proud that $Mary_{iF:[SG]}$, **has**_[SG] been to China.

(21)



What is behind the algorithms? One hypothesis is that the difference between how multiple agreement is resolved is driven by the nature of iF and uF. For uF, one can follow the traditional assumption that they come in forms of [SG], [DL], and [PL]. If one singular value is copied onto the target, it forms a singleton set: {[SG]}, which is spelled out as singular. If two singular values are copied onto one target, they form a set of values {[SG],[SG]}. This set is equivalent to the singleton set given the identity of the member values: {[SG],[SG]} → {[SG]}, which is also spelled out as singular, hence (21a).

Since iF agreement is closely connected to reference, one can assume that they are referential indices (following Grosz 2015). The iF of *Mary* and *Sue* would be their referential indices, say iF:[m], iF:[s]. In multiple iF agreement, the target would contain a set of both values: {[m],[s]}. Unlike multiple uF agreement, the result set cannot be reduced to a singleton set. Thus the target would show the plural marking as in (21b). When the two agreement controllers have the same referential index as in (17), the target would contain a set with two identical members e.g. {[m],[m]}, which is reduced to a singleton set {[m]}. See Grosz (2015) for a different implementation. Discussion of further consequences of this approach requires future research and goes beyond the scope of the current paper.⁴

Evidence for the proposed analysis comes from pluralia tantum nouns like *jeans* and *pants*. Pluralia tantum nouns are morphologically plural but refer to singular individuals, e.g. *my favourite jeans* refer to one item of clothing. In (22), *my favourite jeans* and *my favourite pants* refer to the same item of clothing. We might expect then, that since the referent is the same, the verb can be singular, consonant with what we have said above. However, only plural agreement is possible.

(22) [Context: my favorite jeans and my favorite pants are the same item of clothing.]
John thinks that my favorite jeans and Bill believes that my favorite pants are/*is missing.⁵

On the face of it, this seems odd but is also unsurprising, given that the whole point of pluralia tantum nouns is that they always appear in plural and control plural agreement: e.g. **My favorite jeans is missing*. However, Smith (2015) argues, independently of their agreement abilities in RNR contexts, that pluralia tantum nouns are underspecified for an iF

⁴A note on the difference between our proposal and Grosz (2015). Grosz (2015) proposes that number features are referential indices to account for TP RNR, which this paper follows. However, this paper makes explicit that there is a different type of morphological feature that is behind the distributive agreement. This dual feature analysis unifies hybrid and non-hybrid nouns.

⁵Thanks to Thomas McFadden for pointing out this example to us.

for number. The reason is that they are inherently specified for uF :Number, and this blocks them (at least in English) from combining with a NumP in the regular way, where the non-inherent number information would be introduced. Since they lack a iF value for number, it then stands to reason that they cannot participate in semantic agreement, leaving the uF as the only agreement target for the verb. Thus (22) actually provides another piece of evidence for the current proposal. See Grosz (2015) for another argument from quantifiers.

5. Conclusion and future research

This paper takes on the question whether non-hybrid nouns have both semantic and morphological features like hybrid nouns do. Using multi-valuation in TP RNR as a tool, we argue that distributive agreement results from agreeing with morphological features and summative agreement results from agreeing with semantic features. The conclusion is that non-hybrid nouns, at least in some languages including English, have both uF and iF .

This line of research opens up a number of questions. Firstly, if both iF and uF are available on non-hybrid nouns, why is it not the case that all languages take advantages of them in TP RNR (e.g. Slovenian and Dutch)? This question is clearly related to the fact that semantic agreement of potential hybrids varies according to language: that is, not all potential hybrids in a language will control semantic agreement, if semantic agreement is even present in the language at all. Secondly, what other diagnostics are there to tease apart iF and uF besides TP RNR? Thirdly, do other cases of multi-valuation show summative and distributive agreement patterns? This last question is particularly promising given that recent literature has proposed multi-valuation style analyses for local portmanteaux (Gluckman 2016), switch reference (Arregi and Hanink 2018, Clem 2018), and copular constructions (Coon et al. 2017).

References

- Arregi, Karlos, and Emily Hanink. 2018. Switch reference in Washo as multiple subject agreement. In *the Proceedings of NELS 48*, volume 1, 39–48.
- Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Dordrecht: Springer.
- Belk, Zoë, and Ad Neeleman. 2018. Multi-dominance, right-node raising and coordination. Manuscript, University College London.
- Bhatt, Rajesh, and Martin Walkow. 2013. Locating agreement in grammar: An argument from agreement in conjunctions. *Natural Language and Linguistic Theory* 31:951–1013.
- Clem, Emily. 2018. Amahuaca ergative as agreement with multiple heads. *Natural Language & Linguistic Theory*.
- Coon, Jessica, Stefan Keine, and Michael Wagner. 2017. Hierarchy effects in copular constructions: The PCC corner of German. In *the Proceedings of NELS 47*, ed. Andrew Lamont and Katerina Tetzlo, 205–214. Amherst, MA: GLSA.
- Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15:203–224.

- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2006. *Agreement*. Cambridge University Press.
- Gluckman, John. 2016. Decomposing morphological number in local contexts. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Sophie Nickel-Thompson Kelli Finney, Mara Katz, and Lisa Shorten, 146–156. Somerville, MA: Cascadilla Proceedings Project.
- Grosz, Patrick. 2015. Movement and agreement in Right-Node Raising constructions. *Syntax* 18:1–38.
- Kluck, Marlies. 2009. Good neighbors or far friends: Matching and proximity effects in Dutch Right Node Raising. *Groninger Arbeiten zur Germanistischen Linguistik* 48:115–158.
- Larson, Bradley. 2012. A dilemma with accounts of Right Node Raising. *Linguistic Inquiry* 43:143–150.
- Pullum, Geoffrey, and Arnold Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 62:751–773.
- Puškar-Gallien, Zorica. 2019. Resolving polite conflicts in predicate agreement. *Glossa: A journal of general linguistics* 4:33.
- Shen, Zheng. 2017a. Multi-valued Ns and Ts in number concord and agreement. In *University of Pennsylvania working papers in linguistics*, volume 23, article 28.
- Shen, Zheng. 2017b. On multi-valued Ns and Ts in number concord and agreement. In *Proceedings from the 52nd annual meeting of the Chicago Linguistic Society*, 469–483. Chicago, IL: University of Chicago.
- Shen, Zheng. 2018. Feature arithmetics in the nominal domain. Doctoral dissertation, University of Connecticut, Storrs, CT.
- Shen, Zheng. 2019. The multi-valuation agreement hierarchy. *Glossa: A journal of general linguistics* 4:46.
- Smith, Peter W. 2015. Feature mismatches: Consequences for syntax, morphology and semantics. Doctoral dissertation, University of Connecticut, Storrs, CT.
- Smith, Peter W. 2017. The syntax of semantic agreement in English. *Journal of Linguistics* 53:823–863.
- Wechsler, Stephen, and Larisa Zlatič. 2000. A theory of agreement and its application to Serbo-croatian. *Language* 76:799–832.
- Yatabe, Shuichi. 2003. A linearization-based theory of summative agreement in peripheral-node raising constructions. In *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, ed. Jong-Bok Kim and Stephen Wechsler, 391–411. Stanford, CA: CSLI Publications.
- Zaenen, Annie, and Lauri Karttunen. 1984. Morphological non-distinctiveness and coordination. In *Proceedings of Eastern States conference on linguistics*, 309–320.

Zheng Shen and Peter W. Smith

zhengshen@gmail.com, p.smith@em.uni-frankfurt.de