# Morphological Feature Mismatches under Ellipsis: An Eye-tracking Study

Helena Aparicio, Katie Franich, Ming Xiang[1]

The University of Chicago

## 1.      Introduction

One of the central questions in the study of ellipsis is to determine what principles constrain the identity condition defining the relation between the unpronounced material at the ellipsis site and its antecedent (see an overview in van Craenenbroeck & Merchant, 2013; Merchant, 2013 and references therein). There exists ample evidence, including tests on locality, case matching, preposition-stranding, etc., that strongly suggests ellipsis resolution involves computing silent structures at the ellipsis site (see Merchant (2013) for an overview of the existing diagnostics for structures at the ellipsis site). There is also evidence that syntactic parallelism is one of the principles constraining the identity condition on ellipsis. For instance, antecedent-ellipsis morphosyntactic mismatches can result in ungrammaticality. Lasnik (1999) discusses the case of the verb *be* and the auxiliary *have* in English (1).

(1) a. *John was here, and Mary will too <be here>.[2]
    b. *John has left, but Mary shouldn't <have left>.

    c. John was here, and Mary too <was here>.
    d. John has left, but Mary hasn't <left>.                               (Lasnik, 1999)

The ungrammaticality of the examples in (1a) and (1b) contrasts with the grammaticality of the matched control examples in (1c) and (1d), showing that absence of morphosyntactic isomorphism must be the cause of the degraded status of (1a-b).

However, lack of morphosyntactic parallelism has also been claimed to be grammatical in a variety of languages (Depiante & Masullo, 2001; Nunes & Zocca, 2005, 2009; Bobaljik & Zocca, 2011; Merchant, 2014). The perceived generalization seems to be that morphological features, specifically *inflectional* features, are largely irrelevant for ellipsis computation. For instance, Merchant (2014) reports that the Greek example in (2) is well-formed, despite the mismatch in gender features between the ellipsis site and its antecedent.

(2) I    Maria ine ikani,           ala o   Petros dhen ine <ikanos>.
    the Maria is capable.**fem**.sg but the Petros not  is   <capable.**masc**.sg>
    'Petros is capable, but Anna is not.'                           (Merchant, 2014)

Examples such as (2) complicate the empirical picture, and call into question whether the syntactic representation of the antecedent is indeed relevant for ellipsis resolution in such cases. In this paper, we provide empirical evidence from two eye-tracking

[1] Authors are listed alphabetically. All authors contributed equally.
[2] Throughout this paper, angled brackets indicate elided material.

experiments which suggests that, despite their overall high acceptability, elliptical sentences containing conflicting inflectional phi-features *are* more costly to process, compared to elliptical sentences without feature mismatch. Furthermore, compared to elliptical sentences, the feature mismatch penalty is weaker for deep anaphora, and is completely absent for non-elliptical sentences. Our findings suggest that morphological features are relevant for ellipsis computation, and pose challenges for previous accounts of morphological feature mismatch under ellipsis.

## 2.      Morphological Mismatches Under Ellipsis

Grammatical antecedent-ellipsis morphologal mismatches have been described for verbal, nominal and adjectival predicates (Lasnik, 1999; Depiante & Masullo, 2001; Nunes & Zocca, 2009; Bobaljik & Zocca, 2011; Merchant, 2014). Here, we focus on inflectional phi-feature mismatches on adjectival predicates. Example (3a) contains an instance of a gender mismatch under VP ellipsis in Brazilian Portuguese discussed by Nunes & Zocca (2009). In (3a) the antecedent of the elided predicative adjective '*alto*' has masculine agreement morphology. However, the controller of agreement in the second clause '*Maria*' is feminine. A similar example from Spanish is given in (4a). Since Spanish does not allow VP ellipsis, we present a case of stripping. In (4a) the predicate antecedent '*alto*' is singular, but the subject controlling agreement in the second clause '*los profesores*' is plural. The acceptability of examples (3-4a) containing ellipsis contrasts with their overt counterparts in (3-4b), in which failure to establish an agreement relation results in ungrammaticality.

(3) a. O  João é  alt**o**          e  a  **Maria** também é.          Brazilian Portuguese
       the João is tall.masc.sg and the Maria    also    is
       'João is tall and Maria is too.'
    b. Maria é  alt(\*o/a)
       Maria is tall.masc.sg/fem.sg                              (Nunes & Zocca, 2009)

(4) a. El  alumno         es alt**o** y   los profesor**es**         también.              Spanish
       The student.masc.sg is tall and the teachers.masc.pl too.
       'The student is tall and the teachers too.'
    b. \*Los profesores        son alt(\*o/os).
        the  teachers.masc.pl are  tall. masc.sg/masc.pl

Despite the lack of morphological isomorphism on the surface, Nunes & Zocca (2009) propose that syntactic identity nonetheless holds for examples such as (3a), with the assumption that the adjectival predicate enters the derivation uninflected for phi-features. Their analysis of (3a) is illustrated in (5) and (6). Nunes & Zocca propose that inflectional phi-features are generated and hosted in a dedicated functional projection such as AgrP, and are valued in a probe-goal relation throughout the course of the syntactic derivation via Agree (Chomsky, 2001). The derivation of the first conjunct in (3a), which provides the antecedent for ellipsis calculation, proceeds in the two steps shown in (5). First the Agr head values its [-interpretable] phi-features as masculine/singular by probing into the AP (5a). Second, the DP subject '*o João*', base-generated in AP, moves out to [Spec, TP] (5b). Nunes & Zocca remain agnostic as to whether the surface phi-features on the adjective result from syntactic operations, i.e. overt head movement, or through morphological merger of the Agr head and the

adjectival stem in the phonology. For our purpose, the crucial aspect of the analysis is that ellipsis can be licensed if the identity condition is calculated with respect to the bare adjectival stems, as seen in (6).

(5) a. [$_{AgrP}$ Agr $_{masc.sg}$ [$_{AP}$ alt- o João]]
   b. [[O João] é [$_{AgrP}$ Agr$_{masc.sg}$ [$_{AP}$ alt-]]]

(6) a. [[O João] é [$_{AgrP}$ Agr$_{masc.sg}$ **[$_{AP}$ alt-]**]]
   b. [[a Maria] também é [$_{AgrP}$ Agr$_{fem.sg}$ **[$_{AP}$ alt-]**]]

Even though Nunes & Zocca do not discuss this possibility, an alternative account of the mismatches in (3) and (4) is to assume that the identity relation between the antecedent and the ellipsis site is calculated with respect to semantic, not syntactic, representations. For instance, if the elided adjectival predicate is treated as a null proform with no internal syntactic structure (similar to the analysis proposed for VP ellipsis in Dalrymple et al., 1991 and Hardt, 1993), ellipsis resolution would only require a semantic antecedent with a suitable property, and there would be no requirement on the morphosyntactic form of the antecedent. The null proform analysis has also been adopted by Merchant (2014) to account for some of the feature mismatches under nominal ellipsis, although with a different implementation.[3]

Both analytical options above share the assumption that morphological feature mismatch does *not* affect ellipsis computation. In this paper, we present the results of two eye-tracking reading experiments that suggest this assumption is empirically unwarranted: participants in our experiments were highly sensitive to morphological feature mismatch, which led to both degraded acceptability judgments and elevated reading time for elliptical mismatched conditions. Furthermore, the sensitivity to morphological mismatch appeared to be robust for ellipsis, but weaker for the otherwise very similar cases of deep anaphora. Our findings as a whole pose challenges for both accounts described above. More specifically, the contrast between ellipsis and deep anaphora argues against a pure semantic identity account that dismisses any syntactic parallelism between ellipsis and its antecedent. For this reason, accounts such as Nunes & Zocca's, which assume syntactic identity, seem to be more promising. However, as will be discussed below, their analysis makes the wrong predictions for morphological mismatches under stripping such as (4a). Before we move on to describe the current experiments, we introduce the distinction between surface and deep anaphora in the following section.

## 3.    Deep vs. Surface Anaphora
An important distinction made within the literature on ellipsis concerns 'surface' versus 'deep' anaphora (Hankamer & Sag, 1976; Sag & Hankamer, 1984). While the resolution of surface anaphora relies crucially on a linguistically supplied antecedent, deep anaphora can be controlled by a salient discourse antecedent. We can observe this difference in

---

[3] Phi-feature mismatches under nominal ellipsis present a much more complicated empirical landscape, as not all nouns license all the logically possible directions of mismatch. For existing accounts of phi-feature mismatches in the nominal domain, see Depiante & Masullo (2001); Bobaljik & Zocca (2011) and Merchant (2014).

English by contrasting cases of VP ellipsis (VPE), which is considered a form of surface anaphora (7a), with *do-it* anaphora, which is considered a form of deep anaphora (7b).

(7) a. [Sag produces a cleaver and prepares to hack off his left hand]
      Hankamer: #Don't be alarmed, ladies and gentlemen, we've rehearsed this act several times, and he never actually does.
    b. [Same context]
      Hankamer: …He never actually does it.                    (Hankamer & Sag, 1976)

We see that the bracketed discourse antecedent is able to license the statement in (7b), but not the one in (7a). Note that both (7a) and (7b) would be acceptable given an explicit linguistic antecedent (8).

(8)      a. Don't be alarmed, ladies and gentlemen, he always threatens to hack off
            his hand, but he never actually does.
         b. Don't be alarmed, ladies and gentlemen, he always threatens to hack off
            his hand, but he never actually does it.

A number of experimental studies have investigated whether different processes are involved to resolve surface vs. deep anaphora. Particularly relevant for the current purpose are studies that examined whether surface and deep anaphora pattern similarly or differently in the presence of apparent syntactic non-parallelism, e.g. voice mismatch between the antecedent and the anaphora site. For example, Murphy (1985) compared surface and deep anaphora (DA) in the following constructions:

(9)      a. Leslie kicked the ball, but Fran wouldn't.            (*Consistent, VPE*)
         b. * The ball was kicked by Leslie, but Fran wouldn't.    (*Inconsistent, VPE*)

(10)     a. Leslie kicked the ball, but Fran wouldn't do it.       (*Consistent, DA*)
         b. The ball was kicked by Leslie, but Fran wouldn't do it.  (*Inconsistent, DA*)

The 'consistent' and 'inconsistent' conditions were formed with matched or mismatched voice features between the antecedent and the anaphora site. Within the VPE conditions, participants judged the mismatch/inconsistent condition to be more unacceptable than the voice match/consistent condition, but this difference was not found for deep anaphora. A similar interaction was also found in Tanenhaus & Carlson (1990), which examined both voice mismatch and syntactic category mismatch, and compared not only VPE vs. *do-it* anaphora, but also VPE vs. null complement anaphora. It is worth noting that the critical interaction emerged only in sensibility judgment measures, but not in response time (similar to Murphy, 1985). In more recent work, Duffield et al. (2009) found that parallelism effects were present for both deep and surface anaphora, but to different degrees. An eye-tracking study by Roberts, et al. (2013) found that lack of parallelism influenced processing of both types of anaphora, but that the time-course of processing differed between the two types of constructions.

Thus, the emerging picture from this literature is that deep anaphora is more sensitive to antecedent-ellipsis mismatch than surface anaphora. Building upon this observation, the current study will compare ellipsis and deep anaphora with regard to morphological mismatch. Consider the ellipsis construction in (11a). Spanish has another

type of construction which is structurally quite similar to (11a), but which involves the neuter particle '*lo*' instead of ellipsis (11b).

(11)    a. Juan es alto          y    María también.
            Juan is tall.masc.sg and  María too
            'Juan is tall and María is too.'
          b. Juan es alto          y    María *lo*      es también.
            John is tall.masc.sg and María it.neut is  too
            'Juan is tall and María is too.'

Most important for the current purpose, the particle '*lo*' demonstrates properties of a deep anaphor. As shown in the example below, the particle '*lo*' can be controlled by a discourse antecedent (12a), but ellipsis cannot (12b).

(12) [The tallest players on the basketball team are getting picked to play on the national team. Juan does not get picked and he yells angrily at the coach:]
        a. Por qué  no  he        sido seleccionado?     Yo **lo**      soy también!
           For what  not  have.1sg  been  selected.masc.sg?  I   **it.neut**  am  too!
           'Why haven't I been selected?  I am (tall) too!'

        b. Por qué  no  he        sido  seleccionado?     #Yo también!
           For what not  have.1sg  been  selected.masc.sg?  I   too!
           'Why haven't I been selected?  I am (tall) too!'

In the experiments below, we examine morphological mismatch under ellipsis in Experiment 1, and the same kind of mismatch is then tested with the particle '*lo*' in Experiment 2.

## 4.      Experiments

Two separate experiments were conducted based on the same design. In the interest of space, we report their design, procedure and results together.

### 4.1    Materials
### 4.1.1   Experiment 1 materials

The target material in Experiment 1 was elliptical structures constructed with a 2x2 design. The first factor is *Match*: the number or gender features on the subject of the second clause either matches or mismatches with the antecedent. Number and gender features are separately tested on two sets of items (see below). The second factor is feature markedness on the subject of the second clause (*SSC*): the subject of the second clause either has a marked feature (plural or feminine) or an unmarked one (singular or masculine). A total of 80 sets of experimental items were constructed. In the first 40 sets of items, the gender feature was manipulated based on Match and SSC, while the number feature was kept constant on all arguments and predicates (i.e. singular). An example is presented in (13). In the second set of 40 items, we manipulated the number feature, while keeping the gender feature constant (i.e. masculine); see example (14).

(13) Elliptical Examples, Gender Set

    a. **Match, Unmarked SSC**

      El        asistente     es organizado    y    el        jefe       < es organizado >    **también.**
      the.m.sg assistant.m.sg is organized.m.sg and the.m.sg boss.m.sg    is organized.m.sg **too.**

    b. **Mismatch, Unmarked SSC**

      La      asistenta     es organizada    y    el       jefe      < es organizado >    **también.**
      the.f.sg assistant.f.sg is organized.f.sg and the.m.sg boss.m.sg    is organized.m.sg **too.**

    c. **Mismatch, Marked SSC**

      El        asistente     es organizado    y    la      jefa      < es organizada >    **también.**
      the.m.sg assistant.m.sg is organized.m.sg and the.f.sg boss.f.sg    is organized.f.sg **too.**

    d. **Match, Marked SSC**

      La      asistenta     es organizada    y    la      jefa      < es organizada >    **también.**
      the.f.sg assistant.f.sg is organized.f.sg and the.f.sg boss.f.sg    is organized.f.sg **too.**


(14) Elliptical Examples, Number Set

    a. **Match, Unmarked SSC**

      El        fugitivo  es peligroso    y    el       preso      < es peligroso >    **también.**
      the.m.sg fugitive.m is dangerous.m.sg and the.m.sg prisoner.m    is dangerous.m.sg **too.**

    b. **Mismatch, Unmarked SSC**

      Los      fugitivos   son peligrosos    y    el       preso      < es peligroso >    **también.**
      the.m.pl fugitives.m are dangerous.m.pl and the.m.sg prisoner.m    is dangerous.m.sg **too.**

    c. **Mismatch, Marked SSC**

      El        fugitivo  es peligroso    y    los     presos      < son peligrosos >    **también.**
      the.m.sg fugitive.sg is dangerous.m.sg and the.m.pl prisoners.m    are dangerous.m.pl **too.**

    d. **Match, Marked SSC**

      Los      fugitivos   son peligrosos    y    los     presos      < son peligrosos >    **también.**
      the.m.pl fugitives.m are dangerous.m.pl and the.m.pl prisoners.m    are dangerous.m.pl **too.**


For every condition above, we also included a non-elliptical version. For example, the non-elliptical counterpart of (13a) above is the grammatical sentence in (15a); for (13c) above, the non-elliptical version is the ungrammatical sentence (15b). These non-elliptical sentences served as fillers, and we didn't analyze them in the results section.

(15) Non-elliptical Fillers, Gender Set

    a. **Match, Unmarked SSC**

      El        asistente     es organizado    y    el       jefe       es organizado    también.
      the.m.sg assistant.m.sg is organized.m.sg and the.m.sg boss.m.sg is organized.m.sg too

    b. **Mismatch, Marked SSC**

      El        asistente     es organizado    y    la      jefa      es organizado    también.
      the.m.sg assistant.m.sg is organized.m.sg and the.f.sg boss.f.sg is organized.m.sg too.


Finally, we also included an additional 50 bi-clausal filler sentences. The filler sentences were all ungrammatical due to subject-verb agreement errors.

## 4.1.2  Experiment 2 materials

Experiment 2 examined deep anaphora sentences containing the neuter particle '*lo*', and was designed similarly to Experiment 1. A total of 80 sets of experimental items and 50 additional bi-clausal fillers were constructed. As in Experiment 1, the number feature was manipulated in the first 40 sets, and the gender feature was manipulated in the second 40

sets. In addition, we included grammatical non-elliptical sentences (i.e. with no agreement errors) as controls. Examples are given in (16-19) below:

(16) Deep Anaphor '*lo*' Examples, Gender Set

    a. **Match, Unmarked SSC**
       El      asistente     es organizado    y    el     jefe    **lo**    **es** también.
       the.m.sg assistant.m.sg is  organized.m.sg and the.m.sg boss.m.sg **it.neut is**  too.

    b. **Mismatch, Unmarked SSC**
       La      asistenta     es organizada    y    el     jefe    **lo**    **es** también.
       the.f.sg assistant.f.sg is  organized.f.sg and the.m.sg boss.m.sg **it.neut is**  too.

    c. **Mismatch, Marked SSC**
       El      asistente     es organizado    y    la     jefa    **lo**    **es** también.
       the.m.sg assistant.m.sg is  organized.m.sg and the.f.sg boss.f.sg **it.neut is**  too.

    d. **Match, Marked SSC**
       La      asistenta     es organizada    y    la     jefa    **lo**    **es** también.
       the.f.sg assistant.f.sg is  organized.f.sg and the.f.sg boss.f.sg **it.neut is**  too.

(17) Deep Anaphor '*lo*' Examples, Number Set

    a. **Match, Unmarked SSC**
       El      fugitivo     es peligroso    y    el     preso    **lo**    **es** también.
       the.m.sg fugitive.m is  dangerous.m.sg and the.m.sg prisoner.m **it.neut is**  too.

    b. **Mismatch, Unmarked SSC**
       Los     fugitivos    son peligrosos    y    el     preso    **lo**    **es** también.
       the.m.pl fugitives.m are  dangerous.m.pl and the.m.sg prisoner.m **it.neut is**  too.

    c. **Mismatch, Marked SSC**
       El      fugitivo     es peligroso    y    los     presos    **lo**    **es** también.
       the.m.sg fugitive.sg is  dangerous.m.sg and the.m.pl prisoners.m **it.neut is**  too.

    d. **Match, Marked SSC**
       Los     fugitivos    son peligrosos    y    los     presos    **lo**    **es** también.
       the.m.pl fugitives.m are  dangerous.m.pl and the.m.pl prisoners.m **it.neut is**  too.

(18) Grammatical Non-Elliptical Examples, Gender Set

    a. **Match, Unmarked SSC**
       El      asistente     es organizado    y    el     jefe    **es organizado**    también.
       the.m.sg assistant.m.sg is  organized.m.sg and the.m.sg boss.m.sg **is  organized.m.sg** too.

    b. **Mismatch, Unmarked SSC**
       La      asistenta     es organizada    y    el     jefe    es organizado    también.
       the.f.sg assistant.f.sg is  organized.f.sg and the.m.sg boss.m.sg **is  organized.m.sg** too.

    c. **Mismatch, Marked SSC**
       El      asistente     es organizado    y    la     jefa    **es organizada**    también.
       the.m.sg assistant.m.sg is  organized.m.sg and the.f.sg boss.f.sg **is  organized.f.sg** too.

    d. **Match, Marked SSC**
       La      asistenta     es organizada    y    la     jefa    **es organizada**    también.
       the.f.sg assistant.f.sg is  organized.f.sg and the.f.sg boss.f.sg **is  organized.f.sg** too.

(19) Grammatical Non-Elliptical Examples, Number Set

    a. **Match, Unmarked SSC**
       El      fugitivo  es peligroso    y    el       preso    **es peligroso**    también.
       the.m.sg fugitive.m is dangerous.m.sg and the.m.sg prisoner.m **is dangerous.m.sg** too.

    b. **Mismatch, Unmarked SSC**
       Los      fugitivos  son peligrosos    y    el      preso    **es peligroso**    también.
       the.m.pl fugitives.m are dangerous.m.pl and the.m.sg prisoner.m **is dangerous.m.sg** too.

    c. **Mismatch, Marked SSC**
       El      fugitivo  es peligroso    y    los     presos    **son peligrosos**    también.
       the.m.sg fugitive.sg is dangerous.m.sg and the.m.pl prisoners.m **are dangerous.m.pl** too.

    d. **Match, Marked SSC**
       Los      fugitivos  son peligrosos    y    los     presos    **son peligrosos**    también.
       the.m.pl fugitives.m are dangerous.m.pl and the.m.pl prisoners.m **are dangerous.m.pl** too.

## 4.2    Participants and procedure

Twenty-eight native speakers of Iberian Spanish (8 males, mean age 32.69) participated in Experiment 1, and 21 participants (18 males, mean age 31.94) in Experiment 2. Subjects' eye-movements were tracked with an EyeLink1000 eye-tracker. The stimuli were displayed in an Acer v173 monitor. All sentences were presented in black font against a white background. Only movements of the right eye were tracked, although viewing was binocular. Before initiating the experiment, participants underwent a standard 9-point calibration procedure. Once the participants were successfully calibrated, they had a series of practice trials to get familiarized with the experiment. In both the practice session and the actual experiment, before each sentence appeared on the screen, subjects had to fixate on a black box that appeared to the left of the screen, coinciding with the beginning of the sentence. The target sentence only appeared once the subject correctly fixated on the box. After reading each sentence, participants moved on to perform a yes-no grammaticality judgment about the sentence they had just read. All the instructions were provided in Spanish.

## 4.3    Data analysis and results
### 4.3.1   Data Analysis

For data analysis, the **critical region** (CR) is defined as the first region in which the morphological feature information on the antecedent becomes relevant for the ellipsis and the '*lo*' conditions, and where local agreement is computed for the non-elliptical full structure. In the elliptical conditions, the CR was the sentence-final word '*también*', plus the period that followed. The CR for the deep anaphora '*lo*' sentences was identified as the particle '*lo*' plus the copula, and the CR for the grammatical non-elliptical constructions was the copula '*es*' plus the following adjective. All CRs are bolded in examples (13-19). Data for three subjects (two from the first experiment and one from the second) were removed from analysis due to substantial tracking loss.

    We analyzed four fixation types including First Fixation duration (FF), First Pass duration (FP), Regression Path duration (RP), and Total duration (TT). FF is measured as the duration of the first fixation in any given region. FP is calculated by summing up all fixations within a region from the first time the eye enters the region from the left until the time the eye leaves, whether to the left or the right. RP includes the First Pass duration for a given region, plus any subsequent fixations that the eye makes to the left of that region (when re-reading previous regions of a sentence), until the eye exits the region

for the first time to the right. Both FP and RP measure processing complexity during the early stages of processing. Finally, TT estimates the overall processing complexity for a region at a relatively late stage of processing. This is done by summing up all fixations within a region, regardless of whether the eye has exited this region in either direction.

Fixation data were analyzed using mixed effect linear regression models fitted in the *lmer* package for *R* statistical programming software (Bates et al., 2014). For grammaticality judgments, a mixed effect logistic regression model was run. All models included Match (2 levels: Match or Mismatch), Subject feature of Second Clause (SSC: 2 levels: Masculine or Feminine), and Feature (2 levels: Gender or Number). For all models, subject and item were included as random effects. Following Barr et al. (2013), maximal models were constructed using all possible random slopes and intercepts, wherever these models were able to converge.

### 4.3.2  Results
*Grammaticality Judgments*
Results from grammaticality judgments indicate that, overall, all items were highly acceptable (≥80% 'yes' responses across all conditions, Figure 1 below). There is nonetheless a robust main effect of Match for the elliptical sentences (mixed effects logistic model, p < 0.001), indicating that participants did prefer sentences with matched phi-features. A similar effect of Match was found for particle-*lo* sentences (p < 0.001), but no effect was found for non-elliptical sentences (p>.8). For Ellipsis conditions, there was also a marginal three-way interaction between Match, SSC and Feature (p<0.07), driven by the fact that sentences in conditions with mismatched gender features where the subject of the second clause was masculine (i.e. *La asistenta es organizada y el jefe también.* 'The.fem.sg assistant.fem.sg is organized.fem.sg and the.masc.sg boss.masc.sg too.') were considered worse than other mismatch conditions. No interactions were found for the particle-*lo* and the non-elliptical conditions.
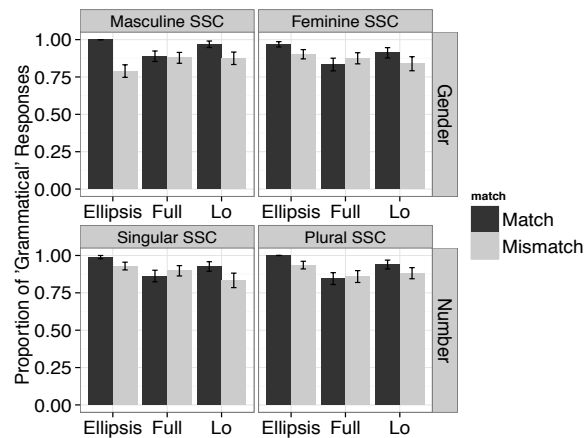


Figure 1: Grammaticality Judgments

*Reading Time On the CR*
Since the most robust effects appeared in the RP and TT measures, in the interest of space, we only discuss results from these two measures below.

*Regression Path time*
For elliptical conditions, results for RP showed a significant main effect of Match ($\beta = 0.16$, t = 5.09, p < 0.0001) (Figure 2), indicating that sentences with mismatched phi-features between antecedent and ellipsis site incurred longer reading times than did sentences with matched features. Neither the interaction between Match and SSC nor the interaction between Match and Feature was found to be significant, suggesting mismatches were just as costly for number features as for gender features, and regardless

of whether the marked (feminine/plural) or the unmarked (masculine/singular) feature value appeared on the subject of the second clause. Critically, no effect of Match (nor any other effect) was found for particle-*lo* conditions, nor was there an effect found for the non-elliptical sentences.
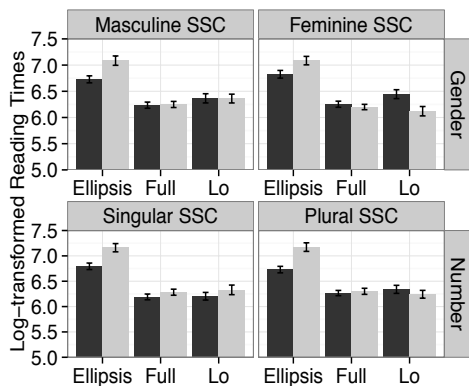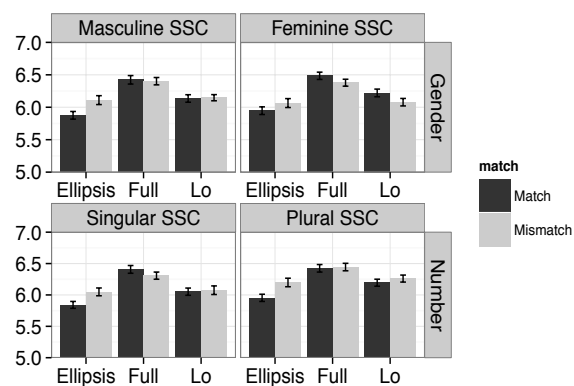

Figure 2: Regression Path Times, CR


Figure 3: Total Time, CR

*Total Time*
Similar to the RP measure, results from the TT measure indicated a main effect of Match for the elliptical conditions ($\beta = 0.08$, t = 3.77, p < 0.0001) (Figure 3), indicating once more that sentences with mismatched phi-features incurred higher processing costs than did sentences with matched features. Again, no significant effects for the interaction between Match and SSC or Match and Feature were found.

*Reading Time On the CR+1*
For both particle-*lo* and non-elliptical constructions, we also examined the region after the CR. This is not possible for the elliptical condition due to the fact that the CR is sentence-final.
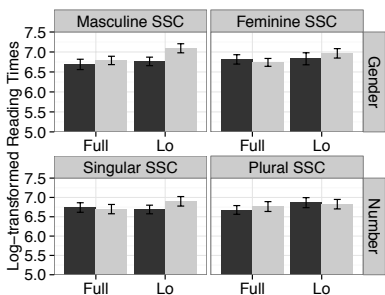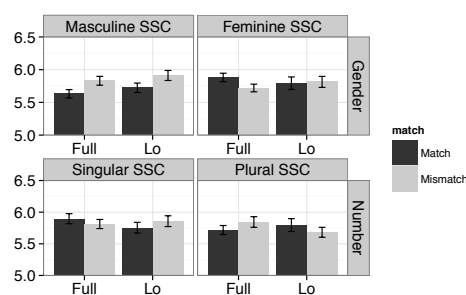

Figure 4: Regression Path Times, CR+1


Figure 5: Total Time, CR+1

*Regression Path time*
For RP, a significant effect of Match was found in the post-critical region for particle-*lo* constructions ($\beta = 0.08$, t = 2.06, p < 0.05) (Figure 4). No effects were found for this fixation type for the non-elliptical constructions.

*Total Time*
No significant effects for the Total Time measure were found at the post-critical region for particle-*lo* constructions. For non-elliptical constructions, there is no main effect of Match, but there is a three-way interaction between Match, SSC and Feature ($\beta = -0.07$,

t = –3.32, p < 0.001) (Figure 5), stemming from the fact that for the gender set of sentences, mismatched features had different effect depending on whether the SSC is masculine or feminine.

## 5.    Discussion

Our results show that, contra previous claims, ellipsis resolution is sensitive to morphological feature identity. This conclusion is supported by the increased reading time at the critical region (CR) for the mismatched ellipsis conditions compared to the matched ones. Furthermore, within the ellipsis conditions, mismatched sentences were judged less acceptable compared to matched ones. As mentioned earlier (section 2), there are two potential approaches to deal with the apparent lack of morphological identity under ellipsis—the first one assumes no morphological mismatch when ellipsis identity is calculated (e.g. Nunes & Zocca, 2009), whereas the second one assumes pure semantic identity (e.g. Dalrymple et al., 1991 and Hardt, 1993).  Our findings raise some important questions and challenges for both approaches, which we discuss in more detail below.

First of all, the current data is not compatible with any account that states that semantic identity is the only identity condition held between the antecedent and the ellipsis site. The observed difference between the particle '*lo*' conditions (deep-anaphor) and the ellipsis conditions (surface anaphor) is informative in this regard. At the critical region, which is the first time point at which the antecedent information is retrieved for anaphora resolution, morphological mismatch only affected the reading time on the ellipsis conditions, but not the '*lo*' conditions. It was not until the next region (CR+1) that morphological mismatch incurred a processing penalty for the '*lo*' conditions as well.[4] For both ellipsis and '*lo*' conditions, morphological mismatch led to decreased acceptability judgments. The different time course between the ellipsis and the '*lo*' conditions with respect to their sensitivity to morphological mismatch is in line with previous findings on voice mismatch from Roberts et al. (2013).[5] Our findings point at the existence of two different mechanisms underlying anaphor resolution for deep vs. surface anaphora, as has also been supported by previous work (Murphy, 1985; Tanenhaus & Carlson, 1990; Duffield et al., 2009). A purely semantic account of ellipsis resolution would not have predicted this difference.

The processing penalty associated with morphological mismatch in the ellipsis conditions lends support to the notion of morphological identity, but it is not completely compatible with the implementation in Nunes & Zocca's (2009) proposal. The analysis in Nunes & Zocca (2009) assumes that apparent morphological mismatches are all grammatical (at least for the adjectival predicates discussed in this paper), since there is no real morphological mismatch in the syntax. However, the current study clearly found acceptability degradation, in addition to elevated online reading time, due to morphological mismatch. In principle, increased reading time and decreased acceptability do not necessarily entail ungrammaticality of mismatched sentences, since these effects could also appear as a result of parsing—grammatical sentences that involve additional parsing steps are known to evoke additional cost. It is therefore possible to maintain that

---

[4] A potential worry is that, for elliptical conditions, the critical region also corresponds to the end of the sentence, which could have triggered additional processing strategies for the elliptical conditions.

[5] Note that our results differed somewhat from Roberts et al. (2013) in that we found both an early (RP) and late (TT) effect of Match for elliptical sentences at the critical region, whereas only late effects of parallelism were found for ellipsis in their study.

these sentences with morphological mismatches are grammatical, and explain the current results in terms of the particular operations the parser must perform during the processing of feature mismatch at the ellipsis site. This approach has been adopted in previous studies to explain the variable acceptability judgments on voice mismatch under ellipsis (Kim et al., 2011).

However, if the current results were to be explained as processing rather than grammatical effects, one would also need to provide a grammatical analysis to derive the well-formedness of elliptical sentences with feature mismatches. Extending Nunes & Zocca's analysis to Spanish mismatch under stripping turns out to be a non-trivial matter. The account proposed by Nunes & Zocca (2009), was developed to handle morphological mismatch under post-auxiliary ellipsis in Brazilian Portuguese. However, Brazilian Portuguese and Spanish differ in that the latter does not allow post-auxiliary ellipsis, as shown in (20).[6]

(20) Juan es alto            y   María también (*es).
      Juan is tall.masc.sg  and María   also        is
     'Juan is tall and María too.'

(21a) contains a direct extrapolation of Nunes & Zocca's analysis to the Spanish facts. As can be observed in (21a), by eliding a low ellipsis site only including the AP, we can correctly account for the availability of phi-feature mismatches. However, this account incorrectly predicts that the copula 'es' should be spelled out, contrary to what we see in (20).

(21) a. Low Ellipsis Site



**Antecedent**          **Ellipsis**
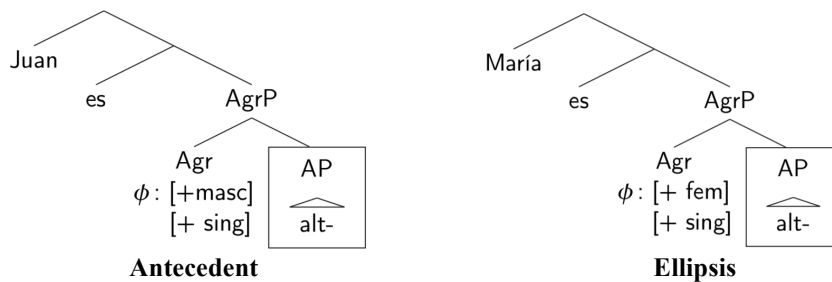
   b. High Ellipsis Site



**Antecedent**          **Ellipsis**

---
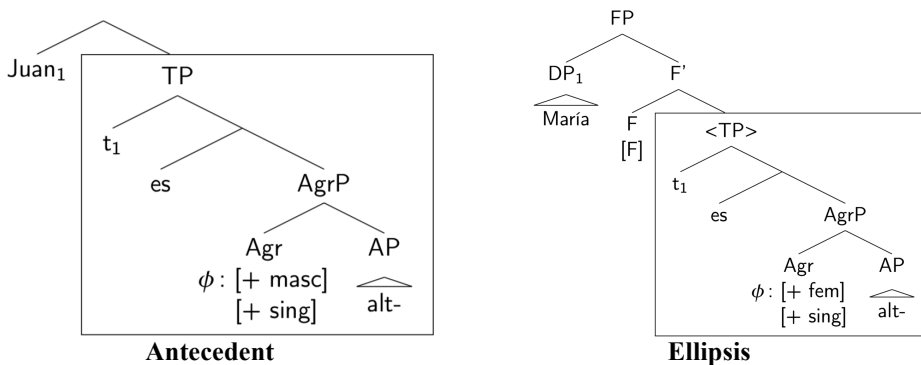
[6] In fact, the same problems arise for Brazilian Portuguese, which also allows mismatches under stripping.

In order to overcome this issue, one could posit a larger ellipsis site containing the copula, as in (21b). This approach corresponds to standard syntactic accounts of stripping, which assume that the deletion operation takes place at the TP level after the remaining DP, '*María*' in (21b), moves out to a [Spec, FP] position (Merchant, 2003). However, in such an analysis, ellipsis would fail to be licensed, since morphological isomorphism is violated, as the ellipsis site and the antecedent now contain conflicting features.

An alternative approach to the Spanish facts would be to adopt an account that takes morphological mismatches under stripping to be ungrammatical, but rescued by the application of a series of repair operations. This type of analysis has been proposed by Arregui et al. (2006) to account for the variability in voice-mismatch acceptability (i.e. the 'recycling hypothesis'). The recycling hypothesis assumes that the antecedent structure is copied into the ellipsis site under structural parallelism. In the absence of structural parallelism, the antecedent can be altered under a series of repair rules to create an antecedent of the right shape. One potential challenge for this account is to explain why morphological mismatches tested in the current study are rated as highly acceptable ($\geq$80% acceptance rate across the board), despite taken to be ungrammatical. More importantly, the specific principles invoked for rescuing and repairing phi-feature mismatches would have to be explicitly defined. For example, one may assume a feature deletion step that erases the incompatible feature on the antecedent before copying it into the ellipsis site.

## 6.      Conclusion

To summarize, we have provided empirical evidence to show that ellipsis resolution is sensitive to morphological mismatches between the antecedent and the ellipsis site. Sensitivity to mismatch suggests that inflectional phi-feature mismatches can no longer be assumed to be irrelevant for ellipsis calculation. Furthermore, the contrast between ellipsis and the deep anaphora '*lo*' shows that a purely semantic formulation of the identity condition will not suffice to capture the patterns of Match effects displayed by the ellipsis conditions; the greater sensibility of elliptical conditions to morphological feature mismatch is not unexpected if ellipsis resolution is sensitive to the syntactic representation of the antecedent. However, further research will be required in order to determine what the best grammatical analysis of morphological mismatches under stripping is, and what grammatical and processing factors contribute to the processing cost observed in the current study, in both online and offline measures.

**References**

Arregui, A., Clifton, C., Jr., Frazier, L., & Moulton, K. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language 55*(2): 232–246.

Baayen, H., Davidson, D., & Bates, D. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language 59*: 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language 68*: 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4. [R package].

Bobaljik, J., & Zocca, C. 2011. Gender markedness: the anatomy of a counter-example. *Morphology 21*(2): 141–166.

van Craenenbroeck, J., & Merchant, J. 2013. Ellipsis phenomena. In *The Cambridge Handbook of Generative Syntax*, ed. M. den Dikken,701–745. Cambridge: Cambridge University Press.

Dalrymple, M., Shieber, S., & Pereira, F. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy 14*: 399–452.

Depiante, M., & P. Masullo. 2001. Género y Número en la Elipsis Nominal: Consecuencias para la Hipótesis Lexicalista. Paper presented at *I Encuentro de Gramática Generativa*, Universidad Nacional del Comahue.

Duffield, N., & Matsuo, A. 2009. Native-speakers' vs. L2 learners sensitivity to parallelism in VP-Ellipsis. *Studies in Second Language Acquisition 31*: 1–31.

Hankamer, J., & Sag, I. 1976. Deep and surface anaphora. *Linguistic Inquiry 7*: 391–428.

Hardt, D. 1993. Verb phrase ellipsis: Form, meaning and processing. Unpublished doctoral dissertation, University of Pennsylvania.

Kim, C., Kobele, G., Runner, J., & Hale, J.T. 2011. The acceptability cline in VP ellipsis. *Syntax 14*(4): 318–354.

Lasnik, H. 1999. Verbal morphology: *Syntactic Structures* meets the Minimalist Program. In *Minimalist analysis*, eds. H. Lasnik, 97–119. Malden, Mass.: Blackwell.

Masullo, P.J., & Depiante, M. 2004. Variable vs. Intrinsic Features in Spanish Nominal Ellipsis. Ms., University of Pittsburgh.

Merchant, J. 2003.  Remarks on stripping. Ms., University of Chicago.

Merchant, J. 2013. Diagnosing ellipsis. In *Diagnosing syntax*, eds. L. Lai-Shen Cheng & N. Corver, 537–542. Oxford: Oxford University Press.

Merchant, J. 2014. Gender mismatches under nominal ellipsis. *Lingua 151*: 9–32.

Murphy, G. L. 1985. Processes of understanding anaphora. *Journal of Memory and Language 24*: 290–303.

Nunes, J., & C. Zocca. 2005. Morphological Identity in Ellipsis. *Leiden Papers in Linguistics 2*: 29–42.

Nunes, J., & Zocca, C. 2009. Lack of morphological identity and ellipsis resolution in Brazilian Portuguese. In *Minimalist Essays on Brazilian Portuguese Syntax*, vol. 142, ed. J. Nunes, 215–236. Amsterdam/Philadelphia: John Benjamins Publishing.

Roberts, L., Matsuo, A., & Duffield, N. 2013. Processing VP-ellipsis and VP-anaphora with structurally parallel and nonparallel antecedents: An eye-tracking study. *Language and Cognitive Processes*, *28*(1-2): 29–47.

Sag, I., & Hankamer, J. 1984. Towards a theory of anaphoric processing. *Linguistics and Philosophy 7*: 325–345.

Tanenhaus, M. K., & Carlson, G. N. 1990. Comprehension of deep and surface verb phrase anaphors.  *Language and Cognitive Processes 5*(4): 257–280.